AGARD-AR-356

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT
BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

AGARD ADVISORY REPORT 356

A Designer's Guide to Human Performance Modelling

(la Modélisation des performances humaines: manuel du concepteur)

This Advisory Report has been prepared at the request of the Aerospace Medical Panel.

NORTH ATLANTIC TREATY ORGANIZATION

Published December 1998

Distribution and Availability on Back Cover

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

BP 25, 7 RUE ANCELLE, F-92201 NEUILLY-SUR-SEINE CEDEX, FRANCE

AGARD ADVISORY REPORT 356

A Designer's Guide to Human Performance Modelling

(la Modélisation des performances humaines: manuel du concepteur)

This Advisory Report has been prepared at the request of the Aerospace Medical Panel.

North Atlantic Treaty Organization

Organisation du Traité de l'Atlantique Nord

The Mission of AGARD*

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

— Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;

— Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);

— Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;

— Improving the co-operation among member nations in aerospace research and development;

— Exchange of scientific and technical information;

— Providing assistance to member nations for the purpose of increasing their scientific and technical potential;

— Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

* AGARD merged with the Defence Research Group of NATO (DRG) on 1 January 1998 to form the Research and Technology Organization (RTO) of NATO. However, both AGARD and DRG will continue to issue publications under their own names in respect of work performed in 1997.

The content of this publication has been reproduced
directly from material supplied by AGARD or the authors.

Printed on recycled paper

Published December 1998

ISBN 92-836-1077-6

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

A Designer's Guide to Human Performance Modelling

(AGARD AR-356)

Executive Summary

The Human Performance Modelling Working Group (WG-22) was convened in 1995 as a joint effort between the Flight Vehicle Panel (FVP) and the Aerospace Medical Panel (AMP) of the Advisory Group for Aerospace Research and Development (AGARD).

The overall objective of the Working Group was to provide advice to system designers in the selection, application and use of human performance models (HPMs). Over the past few decades, tools and techniques for modelling and predicting human performance in complex systems have evolved and matured. Some of these tools are now ready to be integrated into the engineering process. A significant amount of work in evaluating and categorising different types of models has been carried out previously. It is clear from these reviews that available models vary considerably in focus and capability and that their widespread use in the design process is relatively new. Thus, the Working Group members felt that system designers would be more likely to benefit from guidance in selecting and applying the appropriate model(s) than from simply reading another catalogue of available models.

The working group achieved its goal by investigating the state of the art in performance modelling, exploring different methods of integrating HPMs into the system design process, demonstrating typical uses of different classes of models through case studies, and developing a prototype expert system. The Human Operator Modelling Expert Review (HOMER) was developed using a representative set of models that included control, sensory, anthropometric, workload, human error and task network models. The logic upon which HOMER was based (e.g., model selection criteria) is included in the report along with a comprehensive taxonomy of model types to ensure that system designers consider all of the relevant factors associated with the selection and use of the models.

The topics addressed in the Report include:

- The Uses and application of HPMs within the design life cycle
- A taxonomy of models
- An assessment of model capabilities and their limitations
- Commercial issues associated with the development and use of HPMs
- Integration of HPMs into the Systems Engineering process
- Validation Issues
- Usability Issues
- The use of an expert system as a means to select an appropriate model

The outcome of Working Group 22 is as follows:

- A prototype expert system (HOMER) for selecting HPMs
- Recommendations to system designers in the use and application of HPMs
- Recommendations to model developers
- Examples of current uses in terms of case study walkthroughs

La modélisation des performances humaines : manuel du concepteur

(AGARD AR-356)

Synthèse

Le groupe de travail No. 22 sur la modélisation des performances humaines a été créé en 1995 à l'initiative conjointe des Panels de la conception intégrée des véhicules aérospatiaux (FVP) et de la médecine aérospatiale (AMP) du Groupe consultatif sur la recherche et les réalisations aérospatiales (AGARD).

L'objectif principal du groupe a été de fournir des conseils aux concepteurs systèmes concernant le choix, l'application et la mise en œuvre des modèles de performances humaines (HPM's). Les outils et techniques de modélisation et de prévision des performances humaines dans des systèmes complexes ont évolué au cours des dernières décennies et ont atteint, aujourd'hui, un certain niveau de maturité. Certains de ces outils sont maintenant prêts à être intégrés au processus de conception. Des progrès non négligeables ont déjà été réalisés dans l'évaluation et le classement par catégorie des différents types de modèles. Il apparaît clairement des résultats de ces travaux que les modèles actuels peuvent varier considérablement du point de vue de leur précision et de leurs capacités, et que l'intégration généralisée de ces modèles au processus de conception est un phénomène relativement récent. Ainsi, les membres du groupe de travail étaient de l'avis que les concepteurs systèmes tireraient plus de profit de conseils en matière de sélection et d'application de modèles appropriés, que de la simple lecture d'un nouveau catalogue de modèles disponibles.

Le groupe de travail a atteint son objectif en établissant l'état actuel des connaissances dans le domaine de la modélisation des performances, en examinant les différentes méthodes permettant l'intégration des HPM au processus de conception, en démontrant les applications caractéristiques des différentes catégories de modèles à l'aide de cas d'études, et en développant un prototype de système expert. Le système expert de modélisation de l'opérateur humain (HOMER) a été développé à l'aide d'un jeu représentatif de modèles sensoriels, anthropométriques, de contrôle, de charge de travail, d'erreur humaine et de réseaux de tâches. Une description de la logique dont HOMER s'inspire (les critères de choix des modèles par exemple), est incluse dans le rapport, avec la taxonomie complète des types de modèles afin que les concepteurs systèmes puissent prendre en considération l'ensemble des facteurs associés au choix et à la mise en œuvre des modèles.

Les sujets examinés dans ce rapport comprennent :

- L'emploi et les applications des HPM dans le cycle de conception.
- Une taxonomie des modèles
- Une évaluation des capacités des modèles et de leurs limitations
- Les aspects commerciaux du développement et de la mise en œuvre des HPM
- L'intégration des HPM au sein du processus de l'ingénierie des systèmes
- La validation
- L'exploitabilité
- L'intérêt d'un système expert pour le choix d'un modèle approprié.

Les résultats des travaux du groupe de travail No.22 se résument comme suit :

- Un prototype de système expert (HOMER) pour le choix des HPM
- Des recommandations à l'intention des concepteurs systèmes concernant l'emploi et les applications des HPM
- Des recommandations à l'intention des développeurs de modèles
- Des exemples d'applications courantes sous forme de cas d'études

Contents

Preface

The application of human performance modelling within the early phases of the design life-cycle can play an important part in optimising the allocation of function and interaction between the human and machine. It will enable human limitations to be considered before commitment to complex system design solutions that are costly to modify at later stages of the design life-cycle. Working Group 22 was convened in 1995 to address the issues associated with using and developing human performance models. The principal target audience for the Report and its related expert system (HOMER) is all military and industrial organisations involved in the specification, procurement, design, qualification and certification of military systems where the human contribution impacts on mission effectiveness. Model developers within commercial and research organisations should also benefit from the chapters that deal with model limitations and implementation issues.

It is important to recognise previous approaches to performance modelling to ensure that the proposed output is not duplicating work that has been carried out already. During the inaugural meeting of the Working Group (WG) in Belgium (April 1995), the various activities known to the working group were identified. These included:

- Defence Research Group (DRG) Panel 8, Research Study Group (RSG)-9 1982-1990 (Ref 1)
- AGARD Aerospace Medical Panel (AMP) Working Group 12: Human Performance Assessment Methods 1987-1989 (Ref 2)
- National Research Council 1986-1989 (Ref 3)
- The Technical Co-ordination Programme (TTCP) Human Factors in Aircraft Environments (UTCP-7) 1992 - Current
- Air Standardisation Co-ordinating Committee (ASCC) WP61 1993-1995
- SAE Human Modelling System Users Survey 1994 - Current

It was apparent that there is considerable variation in the capabilities of the models and tools and the WG agreed that the system designer would need guidance in the selection of the appropriate model. The approach taken was to establish a set of attributes that characterised all types of models and to determine the extent to which each model or tool satisfied the attribute constraint. A set of thirteen models, which were representative of the type and range of performance models, was chosen to carry out this classification activity.

At the second meeting in the US (October 1995) a review was conducted of the models under evaluation. The model characteristics were further developed into a form that would be compatible with an expert system. The application of models within the system design process was considered in more detail, particularly their potential use within the qualification process. The WG was given a demonstration of the MIDAS integrated modelling environment at NASA-Ames Research Center.

The third meeting was in the Czech Republic in April 1996. A demonstration of most of the tools under evaluation was provided to enable the WG to achieve a greater appreciation of their capabilities. The WG then focused on developing the expert system (Human Operator Modelling Expert Review [HOMER]) and examined all the different criteria a system designer might consider important in terms of his problem domain, his knowledge and experience, the available resources, and so on. A set of 22 questions was drawn up and a score for each of the thirteen models against each of the 80 possible answers was allocated, based upon the capability of the model to answer the specific question. The questions intended to discriminate among competing models were also weighted in terms of their importance to the system designer (e.g., budget) so that inappropriate models/tools are not offered. The WG agreed that another form of 'educating' designers in the use of models was by means of walkthroughs that would provide graphical representations of the use of the tools to solve a specific problem. In this way the system designer could gain a greater insight into the complexity or otherwise of the process by which

1. A Directory of Human Performance Models for System Design (1991)
   DRG AC/243 Panel 8 TR/1
2. Human Performance Assessment Methods (1989)
   AGARDograph 308
3. Human Performance Models for Computer-Aided Engineering (1989)
   NRC Elkind, Card, Hockberg, and Messick-Huey

the required measure of human performance could be obtained. Representative case studies were selected for inclusion in the Report and the overall format of the final report was agreed at this meeting.

The fourth meeting was held in the UK in October 1996. The prototype expert system containing about 80 rules was reviewed by the group. The questions and answers were further developed and the weighting system was refined to ensure that the system dealt with 'show-stoppers' to prevent the system designer being offered unsuitable models. A set of candidate models for inclusion in the final version was identified and a questionnaire was designed to send out to all model developers. The WG was given a demonstration of IPME at the DERA Centre for Human Sciences, Farnborough.

The fifth meeting was held in the Netherlands in April 1997. The meeting concentrated upon completing the chapters of the Report and carrying out further validation of HOMER.

The final meeting was held in the US in October 1997, and included a final review of the Report and HOMER. The commercial aspects associated with maintenance of the expert system is beyond the scope of the Working Group but Micro Analysis and Design is currently hosting the expert system at its web site (WWW.MAAD.COM/AGARD).

Human performance modelling is a key technology that is needed to enable the cost-effective procurement of military systems. Therefore it is important to ensure that the potential users are aware of all the considerations that should be taken into account in the application and use of performance models when applied to their problem domains. The development of the selection criteria and their associated weightings formed an important output of the working group.

Acknowledgements

The authors wish to thank the following personnel who contributed in the provision of presentations of the human performance modelling environments in the UK and US:

Dr. Andy Belyavin	DERA, CHS, Farnborough, Hants, UK
Mr Barry Smith	NASA-Ames Research Center, Moffett Field, California, USA
Mr Sherman Tyler	NASA-Ames Research Center, Moffett Field, California, USA
Mr Jay Shively	NASA-Ames Research Center, Moffett Field, California, USA
Mr Adolph Atencio	NASA-Ames Research Center, Moffett Field, California, USA
Mr Christian Neukom	NASA-Ames Research Center, Moffett Field, California, USA

List Of Acronyms

ANSI	American National Standards Institution
ATM	Air Traffic Management
BAe	British Aerospace
CAD	Computer Aided Design
FAIT	Function Allocation Issues and Trade-offs
GOMS	Goals, Operators, Methods, and Selection rules
HOS	Human Operator Simulator
HOMER	Human Operator Modelling Expert Review
HRA	Human Reliability Analysis
HPM(s)	Human performance model(s)
IPME	Integrated Performance Modelling Environment
MHP	Model Human Processor
MIDAS	Man-Machine Integrated Design and Analysis System
MS HOS	Micro Saint Human Operator Simulator
NASA	National Aeronautical and Space Administration
OCM	Optimal Control Model
POP	Prediction of Operator Performance
PROCRU	Procedure Oriented Crew Model
PSF	Performance Shaping Factor
PUMA	Performance and Usability Modelling in ATM
SA	Situation Awareness
SAFOR	Safe All-Weather Operations for Rotorcraft
SRC	Sowerby Research Centre (part of BAe)
TAWL/TOSS	Task Analysis WorkLoad/TAWL Operator Simulation System
V&V	Verification and Validation
WAS	Aircrew Workload Assessment System
WC Fielde	Workload Consultant for Field Evaluation

Aerospace Medical Panel

Chairman: Lt Col E. Alnaes
Oslo Military Clinic
Oslo Mil/Akershus
N-0015 Oslo, Norway

Deputy Chairman: Dr. A.J.F. MacMillan
RAF SAM
Farnborough
Hants, GU14 6SZ, UK

Working Group 22 Members

Mrs S Hart (Co-Chair)
NASA-Ames Research Center
Moffett Field, CA 94035, USA

Ms J. M. Davies (Co-Chair)
British Aerospace Defence Ltd
Military Aircraft Division
York House
Farnborough Aerospace Centre
Farnborough, Hants, GU14 6SZ, UK

Dr H. Bohnen
National Aerospace Laboratory
PO Box 90502
1006 BM Amsterdam, The Netherlands

Ms G. Burrett
GEC Marconi Research Centre
West Hanningfield Road
Great Baddow
Chelmsford, Essex CM2 8HN, UK

Dr K Corker
NASA-Ames Research Center
Moffett Field, CA 94035, USA

Prof E Farmer
DERA, Centre for Human Sciences
Farnborough, Hants GU14 0LX, UK

Mr P Day
Rokey Manor Research Ltd
Roke Manor, Hants S051 UZN, UK

Dr P Emmerson
Sowerby Research Centre
PO Box 5 - Filton
British Aerospace Defence Ltd
Bristol, BS12 7QW, UK

Dr R Laughery Jr
Micro Analysis & Design
4900 Pearl East Circle
Boulder, Colorado 80301, USA

Dr C Miller
Honeywell SRC/MN65-2500
3660 Technology Drive
Minneapolis, MN 55418, USA

Mr R Parus
Sextant Avionique
Rue Toussaint Castros
BP 91
33166 Saint Medard en Jalles Cedex, France

Prof N Moray
University of Surrey
Guildford, Surrey GU2 5XH, UK

Other Technical Contributors

Mr G A Ward
ESE Associates
15, Jesse Close
Yateley, Hants, UK

Mr B. Buck,
GEC Marconi Research Centre
West Hanningfield Road
Great Baddow
Chelmsford, Essex CM2 8HN, UK

Dr V Riley
Honeywell SRC/MN65-2500
3660 Technology Drive
Minneapolis, MN 55418, USA

Dr Andy Belyavin
Centre for Human Sciences
DERA
Farnborough, Hants GU14 0LX, UK

Mr Christian Neukom
NASA-Ames Research Center
Moffett Field, CA 94035, USA

Panel Executive

Major R. Poisson, CAF
AGARD-OTAN - AMP Executive
7, rue Ancelle - 92200 Neuilly-sur-Seine- France
Telephone: 33(01)55 61 22 60/62
Telex: 610176 (France)
Telefax: 33 (01) 55 61 22 99

1. INTRODUCTION

Human performance is often a high-risk element in the operational effectiveness of complex systems. For example, more than two thirds of all aircraft accidents continue to be attributed to pilot error, and the human element cannot be ignored. The traditional design process has placed a disproportionate focus on the technical performance of equipment, with little regard for the human component. In fact, even equipment design is narrowly focused on the functional performance of the equipment, rather than its actual contribution to overall mission effectiveness. A greater emphasis on mission performance requirements would help enable more accurate trade-offs to be made among sub-systems, allow identification of critical success criteria, and facilitate more effective evaluation of fitness for purpose. This perspective would enable the integration of Human Performance Models (HPMs) into the design process.

In the past, it has been difficult to integrate HPMs into system performance models, because of the complexity of human behaviour and the lack of computational power to address the variability in human performance. Traditionally, the techniques that have been used to examine human performance issues have been largely manual and laborious in nature. However, modern tools and methods facilitate the transfer of this information in a format compatible with other system models. This provides a golden opportunity to ensure that problems associated with human performance are identified early in the design process to prevent costly changes and procurement delays. These integrated models can help give insight into expected human performance and ensure that the technology will support effective collaboration between human and machine to achieve system goals.

A significant amount of work in evaluating and categorising HPMs has been carried out previously. However, the Working Group felt that system designers would benefit more from guidance in selecting and applying models that were most appropriate for their application than from another catalogue of available models. Therefore, the primary objective of the Working Group was to provide advice to system designers in the selection, application and use of HPMs. This was achieved through case studies that demonstrate typical uses of models within the system design process and through the development of a prototype expert system. This system was developed to give designers advice about the relative applicability of different HPMs to their design goals, given practical constraints of time, funds, and staff. The Human Operator Modelling Expert Review (HOMER) currently contains a representative set of control, sensory, perception, anthropometric, biomechanical, workload, human error and task network models. To create HOMER, the Working Group identified a set of questions that the designer should ask during model selection and assigned weights that represented judgements of the relative importance of each. A secondary objective was to provide ideas and insights to the HPM community regarding additional developments that would enhance the application of HPMs to system design.

1.1 Organisation of this Report

This report is organised into six additional chapters and two appendices. These chapters are as follows:

Chapter 2, Applications of HPMs, describes historical applications of HPMs within the system design process. It cites specific examples of how human performance modelling can enhance design effectiveness. This section is intended to stimulate the interest of systems engineers about issues that may be addressed with HPMs.

Chapter 3, Taxonomy of Models, provides a taxonomy of HPMs. There are many different types of HPMs that have evolved over the years; Section 3 categorises them in a way that should be meaningful to systems engineers.

Chapter 4, Model Limitations, identifies known limitations with the current models. This chapter is intended to clarify what current models are not capable of doing well. This is also intended to identify for the modelling community areas where model development would have high potential payoffs.

Chapter 5, Implementation Issues, addresses some of the pragmatic issues associated with the fielding of HPMs. It provides guidance for HPM developers to help ensure that their models will be usable by systems engineers.

Chapter 6, Description of the Expert System, describes the prototype system developed by the Working Group to assist the systems engineer in selecting the appropriate model for a particular application. It describes the rationale and underlying structure of HOMER and offers practical hints about how to use it.

Chapter 7, Recommendations, provides a series of recommendations applicable to systems engineers, users, model creators, and distributors regarding the use, development, and limitations of HPMs.

Appendix A, Case Studies, provides a series of "walkthroughs" that demonstrate how existing HPM tools might be used to study example design problems. They are intended to illustrate different tools and problem domains in concrete terms.

Appendix B, HOMER spreadsheets, details the questions and weightings used in the expert system

2. APPLICATION OF HPMs

This Chapter describes historical applications of HPMs within the system design process. It cites specific examples of how HPMs can enhance design effectiveness and attempts to stimulate the interest of systems engineers about issues with which they may be familiar.

Chapter 6 describes a process for selecting models that will assist in addressing the following issues. The first question the expert system asks the user deals with the very important issue of the application(s) that the model will be expected to address. HPMs have been and can be applied to the following design issues:

1. Operational analysis/operations research
2. Frequency and nature of errors
3. Effects of environmental stressors
4. Requirements development
5. Training requirements
6. Certification
7. Function allocation
8. Automation
9. Crew complement
10. Selection
11. Workload
12. Team interaction
13. Communications
14. Display design and evaluation
15. Control design and evaluation
16. Workspace design
17. Development of procedures

2.1 Operational Analysis/Operations Research

Operational Analysis (OA) is performed to examine the impact of technology on operational effectiveness. Humans play a vital role in the operational effectiveness of both civil and military systems, and key aspects of their performance should be incorporated into OA models. HPMs evaluate the potential impact of factors that are likely to influence human performance and provide data on task times and error rates which can then be incorporated into models of the overall system.

2.2 Frequency and nature of errors

Specialised models are available to predict the types of human error that may be associated with a system design, and the frequency of these errors. Obviously, the large contribution of human error to system failure should be included in safety analyses. In particular, these data should be included in safety and failure mode analyses to ensure that the contribution of human error to system failures is recognised to ensure that the system design is tolerant of likely human errors.

2.3 Effects of environmental stressors

Stressors such as heat, noise and fatigue are known to have particular patterns of effect upon operators' cognitive and physical processes. Some modelling environments take into account the effects of these stressors, thereby providing useful input into safety hazard analyses.

2.4 Requirements development

HPMs can help determine the level of human performance required to meet system performance requirements. This information can be used to perform system level trade-offs and to ensure that sub-systems work together to support effective human performance. In addition, the development of human performance requirements allows traceability of the human component in system design, and facilitates the development of criteria for acceptance tests associated with fitness for purpose.

2.5 Training requirements

The amount of training required is an important design driver. Novel designs, such as new methods of presenting aircraft attitude information, may require considerable training if operators are already experienced in conventional formats, but promise enhanced performance. HPMs can identify areas where an investment in training will have significant human performance benefits and help to assess cost-benefits of training system options.

2.6 Certification

System certification procedures are placing increasing emphasis on human factors issues. More and more customers want evidence that systems will be fit for purpose and that human factors principles have been applied to design. HPMs can provide criteria for assessing total system (both human and machine) performance and provide evidence for the likelihood that acceptable performance will be achieved.

2.7 Function allocation

It is often necessary to determine whether a task would be performed better by a human or by the system. Although simple lists have been developed to indicate the relative strengths of humans, computers, and hardware for performing different tasks, typically it is necessary to model the particular system in question in order to achieve the most effective co-operation between human and machine for a particular function.

Once such a model has been developed, however, analyses of capabilities and availability of human and system resources offered by the HPM can aid in making function allocation decisions based on objective criteria and for a variety of possible circumstances.

2.8 Automation

The advantages of automated systems may be compromised by the disadvantages of removing the operator from the control loop and hence reducing Situational Awareness (SA). It must be ensured that, in the event of failure, the operator will be able to resume manual control. Performance models are available that allow exploration of such issues.

2.9 Crew complement

Modelling can be used to determine the number of operators required for a particular system. This is a critical decision in system design, since it has consequences for the cost of equipment and operator training, the operational effectiveness of the system, and decisions about function allocation and automation.

2.10 Selection

Performance modelling can aid operator selection by indicating special abilities or other operator characteristics required by the equipment or tasks.

2.11 Workload

Several methods of predicting crew workload have been developed. These models are often based upon estimates of the resource(s) demanded by the task (*e.g.*, mental, physical, visual, auditory) and the extent to which different combinations of tasks will interfere with each other when performed concurrently. The underlying models upon which these methods are based differ with respect to assumptions about the number and independence of such resources, combinatorial rules across resources and concurrent tasks, and how instances of 'overload' are handled.

2.12 Team interaction

In multi-operator systems, effective teamwork is essential. Interaction between team members can be modelled during system design. Some HPMs can be used to characterise the flow of information required to perform specific tasks.

2.13 Communications

Methods are available to model communication. For example, recognition rates over noisy channels can be estimated. In addition communication capabilities or limitations can be modelled via speech and auditory modality demands in order to establish designs that permit good communication.

2.14 Display design and evaluation

Modifications to display design, such as the addition of colour coding, may be very costly. The designer must be able to predict the benefits, if any, of such modifications. Operator preference is not a sufficient criterion for any display solution: often, subjective preference is unrelated to performance.

2.15 Control design and evaluation

Modelling can be used to predict the effects on performance of control variables such as lag, control order (position, velocity, acceleration, etc.), and gain. Recently, computational models have been developed that capture the control laws and principles developed over many years, offering them to systems designers in a format that is convenient to use early in design.

2.16 Workspace design

Modelling systems are available to ensure that equipment layout is optimised. They accept and produce 2-D and 3-D renderings of the workspace, compare alternative layouts, and offer feedback about the strengths and weaknesses of each with respect to the position, reach, comfort, viewing angle, etc of human crewmembers having different physical characteristics as well as the logic of control and display placement given the flow of tasks to be performance and information to be processed.

2.17 Development of procedures

The development of procedures for complex systems can be guided by the use of modelling tools.
Procedure timing and the consequences of procedural deviations are two examples of the types of questions that can be studied, as well as associated information flow, display formatting, and information entry options

3. TAXONOMY OF MODELS

This chapter provides a taxonomy of different types of HPMs that have evolved over the years. They are categorised in a way that will make the variety and nature of them meaningful to a systems engineer.

3.1 Introduction

The systematic investigation of human performance began with the attempts by Donders during the 1860s to identify the mental processes underlying the reaction time to simple stimuli. Later in the nineteenth century, Ebbinghaus began a long series of studies of human memory. It was not until World War II, however, that intense interest in the performance of the human operator developed. It was found, for example, that the performance of radar operators quickly declined during a period of duty, and that many aircraft accidents were attributable to pilot error. Recognition of the effects of poor equipment design led to the development of the field of 'ergonomics', in which the psychological, physiological and engineering aspects of man in his working environment were considered; the clear limitations of the human operator interacting with complex systems were addressed by the science of 'cognitive psychology', in which the acquisition, processing and output of information by human operators were investigated systematically.

HPMs can be classified in several ways, depending on the target audience. In general, taxonomy development begins by determining the endpoints of the list, and proceeds by populating the space between the endpoints. Typical endpoints might include:

Prescriptive (Normative) versus Descriptive. Descriptive models indicate how a human is likely to perform a task or predict ideal behaviour, whereas prescriptive models show how the humans should perform if they are able to behave in a rational way that takes into account the information available, the existing constraints, and the risks, rewards and objectives.

Top down versus bottom up. This refers to whether the model is dictated by system goals or human performance capabilities. The former focuses on output (system performance) whereas the latter focuses on the processes leading to performance as well as output.

Single Task (limited scope) versus multitask (comprehensive). This distinguishes modelling used to explore specific elements of a single task e.g using a biomechanical model to assess load lifting limits in detail, as opposed to modelling multi-function tasks like piloting an aircraft using a task network model.

Individual versus team performance. The majority of human performance models are concerned with individual performance. Multi-operator or team performance models deal with the additional levels of complexity imposed by multiple communication interaction paths between operators/machines and operators/operators

For present purposes a taxonomy is proposed based on the theories or tools that underlie the models or serve as a basis for their development. This follows the description suggested by the US National Research Council Panel on Human Performance Modeling (Ref. 3). Shown in the top half of Figure 1, several theoretical approaches to human performance are represented. The lower half of the figure depicts a separate categorisation of models labelled 'pragmatic'. These alternative methods of describing models seem to be required for at least three reasons: (1) some models are data driven and do not require a theoretical basis (*e.g.,* anthropometric models); (2) for other models, there is no underlying theory, even though such a theory would generate substantial improvements in the quality of the predictions (*e.g.,* situational awareness); and (3) other modelling techniques incorporate more than one underlying theory, but are narrowly applied (*e.g.,* Human Reliability Assessment).

The following sections discuss the categories shown in Figure 1.

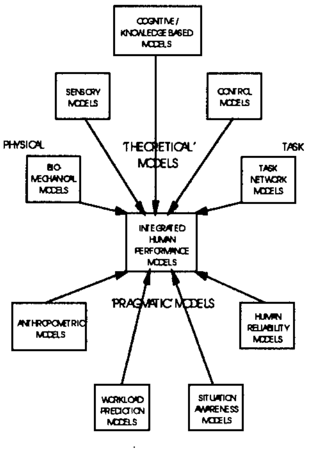


Figure 1 HPM Taxonomy

3.2 Bio-Mechanical Models

In general terms, biomechanics deals with various aspects of the physical movement of the body, using laws of physics and engineering concepts to describe the motion undergone by various body segments and

the forces acting on them. In practice, however, biomechanical models have been used either to predict human materials handling capabilities from calculations that define the body as a mechanical load-bearing device, or have focused on human tolerance limits for vibration and acceleration stress. Many of the latter models are based on existing single-task models with the vibration or acceleration stress represented as a disturbance to visual perception or motor control. Within the class of models that deal with material handling, some attempt to predict lifting capacity, given specific human, task and environmental characteristics, while others use Newtonian mechanics to estimate the stresses imposed on the musculoskeletal system during lifting. Typically, these models are rather restricted, because they assume a limited range of lifting postures and geometries, no mechanical aids, smooth symmetrical lifts and good floor contact.

It should be noted that bio-mechanical models are not the same as anthropometric models, although some anthropometric models do contain aspects of biomechanical limitations. Anthropometric models are used to determine the ability of an operator of a given physical size to work within a given space, to reach specific controls and to see specific displays.

3.3 Information Sensing and Processing Models

This approach describes the human as a processor of sensory and cognitive information. Taking this view, information is passed along a series of sensory channels, starting with the receptors themselves (the eyes, ears etc.), progressing through various temporary holding stores to storage in long-term memory. There are many models that deal with different stages of information processing. For example, some deal exclusively with visual performance, whereas others have been developed to quantify attention, memory, discrete movements and simple reaction times. The only significant attempt to integrate this type of micro model into a model of the whole operator led to the development of the Human Operator Simulator (HOS). This was originated in the late 1960s in the US Navy, as a comprehensive computer modelling tool. The execution of a HOS simulation results in a sequence of operator decisions about what to do at each point in time, based on moment-to-moment mission events and predefined tactics and procedures. A data analysis package that is part of the HOS system provides standard statistical human factors descriptions of events that can be used to support a wide variety of purposes. This is a powerful tool, which later became part of a larger tool set (MS-HOS).

3.4 Knowledge Based /Cognitive Approach

Knowledge-based models of human performance are explanations of how people decide what is to be done to solve a problem. These models provide explicit representations of an operator's decision-making processes, rather than simply assuming that the operator will make a correct decision. This is quite different from the typical goal of HPMs; to predict how accurately or reliably a person can execute a procedure under the assumption that the person knows what is to be done. For example, if a pilot needs to apply more than normal power during take-off, a traditional modelling question would be to determine the distribution of times before the crew noticed the problem. A knowledge-based study might begin by modelling the pilot's decision whether or not to apply more power and then investigate this *decision- making process* under various conditions of visibility, fatigue, workload, etc. In essence, the knowledge-based approach treats human thought as an example of symbol manipulation according to rules that can be modelled with computer programmes, but without assuming that the human brain works like a computer. Cognitive models are one of the fastest growing areas of HPM development, and there is little consensus about exactly what should be modelled or how.

Some models attempt to represent human decision processes, at least in a limited domain, by the use of procedural (if-then) rules. Rule-based approaches try to predict what decision will be made in a given situation. Other models use a goal-driven approach to examine how users will decide what tasks or information to attend to. Another approach is to model the use of information by working memory to support decision making. Still others look at the amount of information that can be processed or the time available to make a decision in order to predict decision-making accuracy.

These models are likely to be most useful in situations in which system performance is limited by what the human operator decides to do, rather than how quickly or accurately it can be done (*e.g.*, for supervisory aspects of performance). They have been applied to problem solving in aircraft systems, although not for quantitative predictions.

3.5 Optimal Control Theory Models

The Optimal Control Model (OCM) deals largely with manual control. The human is viewed as an information processing or control/decision element within a closed loop system (the so-called cybernetic view of the human). In this context, information processing refers to the processes involved in selectively attending to various sensory inputs and using this information, along with the operator's

understanding or model of the system, to arrive at an estimate of the current state of the world. Second, in most models based on this approach, it is assumed that trained operators approximate the characteristics and performance of good or optimal inanimate systems performing the same functions. It is assumed that their performance, and thus that of the overall system, is constrained by inherent human sensory, cognitive and response limitations.

Although apparently dealing with a limited area of performance, OCMs have been applied widely, and the information processing portion of has been extended to tasks other than manual control (*e.g.*, failure detection, monitoring, and decision making.) One of the best known OCM implementations, the Procedure Oriented Crew Model (PROCRU), is a derivative that incorporates the execution of procedures in complex cockpit systems in the context of manual control. In general, OCMs provide data that are analogous to person-in-the-loop simulations, with the additional benefit of providing predictions of the operator's internal states. Although not verifiable through measurement, these predictions can be useful for uncovering or diagnosing system problems. They also provide a variety of outputs related to task demands and operator workload. They seem well suited to highly structured situations with well defined goals, but will be less useful when the operator has flexibility in performing the task. In practice, mathematically 'optimal' solutions are rarely calculated, and sub-optimal solutions tend to be developed that compromise the normative nature of the model and increase the modeller's subjective input. On the other hand, their main limitation is the lack of experimental validation for the overall integrated models. A second, but important practical problem is that use of the models requires a sophisticated mathematical and control theory background.

3.6 Task Network Models

These have developed from operations research, and have been the basis of many early uses of HPMs in complex, practical, real world tasks. A complex system is represented by a network of component processes, each modelled by statistical distributions of completion time and probability of success. The resultant computer programme is run as a Monte Carlo simulation to predict the statistical distributions of measures of overall system performance. An example of a task network from the modelling tool IPME is presented in Figure 2 .
The human is assumed to interact with the environment through a sequence of activities or tasks, which are described by an operator action, an object of that action, and other qualifying or descriptive information (*e.g.*, time to complete the task.) A procedure is a collection of tasks required to accomplish some goal.

A task network is a collection of procedures and tasks that contains hierarchical and sequence information. The human is assumed to be sensitive to global variables such as stress or motivation, and the approach also includes estimates of human and system reliability.

To explore the impact of these variables, moderator functions can shift the time distributions or completion probabilities for all component tasks to be performed by the human, based on the setting of the moderator function. Originally, the output from these models was simply time and accuracy to complete

certain procedures. More recently the output has been expanded to include elements such as mental workload estimates, with loadings for four information processing components (*i.e.*, vision, audition, cognition and perception). Task network models are ideal frameworks in which to embed isolated and independent single-task models of human performance.

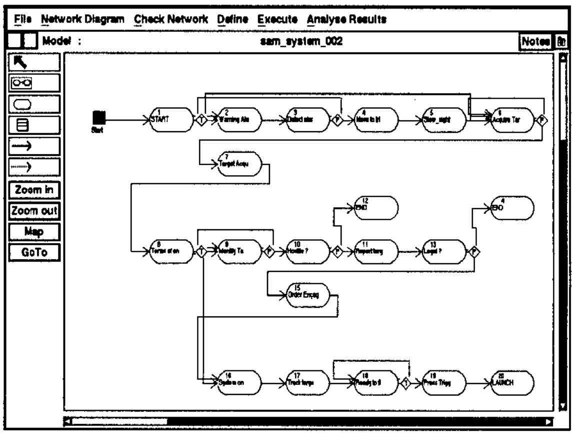


Figure 2. Example of a Task Network from the Modelling Tool IPME

3.7 Anthropometric Models

Anthropometric models are a special form of Computer Aided Design (CAD); they were developed specifically to enable ergonomic design activities to be undertaken in a CAD environment, and their principal feature is a 3-D animated human mannequin. Thus, they focus on the physical relationship between human(s) and their workplace. Anthropometric models are sometimes referred to as Human Modelling Systems or Human Simulation Systems. An example of a display from the anthropometric modelling tool Jack is presented in Figure 3.

Figure 3. Example Display from Jack

Anthropometric models use 3-D animated human mannequins to enable a user to evaluate the ergonomic features of a proposed design solution over the anthropometric range of the target user population. That is, they help determine whether the relationship proposed by a designer between the humans and the controls and displays they use is technically feasible within the constraints of human body dimensions and movement ranges. Traditionally, anthropometric models have been employed to assist in the design and evaluation of complex operator workstations. However, they are equally applicable to issues relating to design for maintainability. Clearly, an important feature of all anthropometric models is their associated database. It must be capable of representing the bodily dimensions of the target user population. Ideally, however, it will be sufficiently flexible to enable different target users and populations to be selected. Current developments of models include a high-resolution figure for use in CAD-based design, and low to medium resolution figures for iconic operator representation in simulations.

This type of model is essential at the start of the design cycle. However, it is also important to re-run the model each time a physical design parameter is modified. When embedded within a simulation, anthropometric models can be used for mission rehearsal.

Typically, this class of model is entirely self-contained. However, it is also possible to import CAD geometry and manipulate the mannequin within this type of environment. The customer may specify the anthropometric range with which a design must comply and must provide the physical dimensions of the workspace in appropriate units. Typical outputs include: (1) Reach envelopes; (2) Eye views; (3) Vision cones; (4) Torque load and comfort during reach; (5) Real-time human-object and object-object collision detection; and (6) Computer 'pictures' of the human in the workplace (which can also be animated.)

3.8 Workload Prediction Models

Workload can be defined as the cost incurred by the human operator in accomplishing the imposed task requirements. This cost reflects the combined effects of the demands imposed by the tasks themselves, the information and equipment provided, the task environment, operator skills and experience, operator strategies, the effort exerted and the emotional response to the situation. It comprises both physical and mental activities. The former can be predicted in the dimensions of time and accuracy, using biomechanical or micro models, but mental workload is rather less straightforward. Essentially there are several theories of how reduced performance under high workload is produced. These address fundamental issues about the nature of human information processing (*i.e.*, serial versus parallel) and will not be explored here; however, they are relevant because they result in different workload models (see below).

It is to be noted that this section deals only with techniques for predicting workload. The measurement of workload (i.e. the subjective or objective calculation of workload on a task that is being or has been performed) is not considered. A guide to measurement techniques for workload is given in the ANSI Guide for Human Performance Measurement (Ref 4).

The general aim of workload prediction techniques is to predict accurately the relationship between task demands and an operator's capacity. The human is assumed to have a number of available channels, containing resources. At issue is whether one can predict the change in performance, given the characteristics of either: (1) the processing on each channel (or task) in isolation or (2) the relationship between channels (tasks).

In practice, the typical objective of a workload analysis is more modest (i.e., to identify peaks in an operator's workload), acknowledging the limited nature of current workload models. These workload peaks are thought to occur as a consequence of an excess of task demands in relation to the operator's available resources. Most models can quantify the factors that contributed to the workload (*i.e.* the individual tasks the operator was engaged in at a particular time, and the effects of those component tasks on workload). Underload conditions are possible as well, although they have received less research. Having identified aspects of a mission that could produce workload peaks, the designer can then examine the individual factors with a view to reducing the demands (*e.g.*, by automating the task or changing the equipment). Other purposes might be to compare the relative

merits of design alternatives or to optimise task-sharing within the team. It should be noted that the process of workload prediction is usually iterative, with the ultimate aim of achieving a design for which workload is at an optimal level. It should also be stressed that, in general, workload prediction models are far less mature than other types of HPM (for example, there is no universal agreement on what constitutes a 'channel'). Indeed, some researchers have begun recently to question the whole concept of multiple resource allocation theory, which is central to many of the current approaches.

The essential input for a workload prediction model is a mission timeline. Ideally, the results of a task analysis that includes the time required for each task should be available as well. The third requirement is for a database of the individual resources demanded by each of the subtasks. Usually, this is formulated in terms of the loading on particular processing channels. Different tools have alternative description of these channels, but typically they will include visual, auditory, cognitive and response/psychomotor. Typical outputs include: (1) Sustained workload (*e.g.*, the average overall workload and how various intensities of sustained workload affect performance); (2) Momentary workload (*e.g.*, the size of workload during peak periods and effect on human performance); (3) Reserve capacity (*e.g.*, the margin of full performance a task requires and an estimate of remaining capacity to perform additional tasks effectively; and (4) Errors (*e.g.*, an estimate of the probability that an error will occur).

Another form of workload analysis is mission timeline analysis, which calculates the ratio between time available and time required to perform the task . A ratio of greater than 1.0 implies that the task cannot be completed, and values between 0.85 and 1.0 are thought to indicate potential workload problems.

3.9 Situational Awareness Models

Situational awareness (SA) can be described simply as "knowing what is going on so that one can figure out what to do" (Adam, 1993) (Ref 5) . In other words, the operator's SA is the sum of the current understanding about the physical environment, system states, own status, and so on. This awareness or knowledge serves as the basis for making critical decisions.

SA is a multi-faceted attribute of human cognition, and this has implications for how it is measured. The purpose of all SA measures is to estimate the operator's level of awareness of the objective situation relative to some ideal level of 'perfect' awareness. It is not feasible, however, to evaluate an operator's awareness of every conceivable item of information at every moment, so SA is selectively sampled. Furthermore, SA measures should be regarded as relative indicators rather than absolute measures.

SA can be assessed with either *objective queries* or *subjective ratings* and may be inferred from other measurements of performance. Whichever technique is used, the aim is to assess the operator's knowledge about: (1) Spatial orientation (*e.g.*, where he is relative to the ground); (2) Positional awareness (*e.g.*, where he has been, where he is going and where he is now); (3) Temporal awareness (*e.g.*, knowledge about events as the task evolves); (4) Automation awareness (*e.g.*, what the system is doing and who is in charge); and (5) Tactical situation awareness (*e.g.*, potential threats).

Objective techniques involve administering a series of queries that 'probe' the operator's knowledge of specific items that are important to the successful completion of the mission. The operator's responses to these queries are then scored against the objective facts of the situation. Alternatively, the speed and accuracy with which an operator responds to specific events might be used as an objective indicator of his SA. Objective assessment gives the most direct measure of SA and can have high validity, if the correct probes, information, or events are introduced. Subjective techniques rely on self reports from the operator during or after a mission or evaluations by an expert observer. Rating scales can be unidimensional or multidimensional. If required, subjective ratings can be taken periodically throughout the course of a mission or after the mission has ended with the memory aid of a video-taped or computer-generated replay. These methods have the potential of providing a task-related profile of SA variation over time. Typical outputs for objective techniques include the proportion of correct responses and the accuracy of numerical responses. Typical outputs for subjective techniques include average ratings and ratings profiles. Objective query techniques require a fairly involved series of information-gathering exercises: (1) Analysis of the tasks to be studied; (2) Expert identification of the information needed to perform each task; (3) Expert evaluation of the priority of each identified information item; (4) Selection of a high-priority subset of information items; (5) Generation of queries based on the selected items; (6) Development of a methodology for presenting queries and recording and scoring responses; and (7) Establishing the correct answers before the study begins.

SA measures can be taken throughout the design cycle, although the types of measures that are most feasible and appropriate will vary with the stage of development. The later in design, and the more integrated and sophisticated the systems under evaluation, the more complex it becomes to administer objective measures. During flight trials, for instance, it is more feasible to obtain subjective measures than objective, performance-based measures.

3.10 Human Reliability Models

The reliability of human-machine interactions refers to the effectiveness with which humans and machines co-operate to accomplish tasks. Neither the human nor machine is assumed to be the sole contributor to reliability. Currently, there are three general groups of approaches to the issue of reliability:

3.10.1 Human error occurs at the level of individual sub-tasks. The different actions a human can perform are distinguished by the accuracy with which they executed, relative to task descriptions. Task-based methods have been developed, based on the notion that human error can be predicted at the level of individual sub-tasks. Most techniques require detailed specifications of tasks before any estimates of interaction problems can be generated. Since the input required is a task analysis, not only is a very well developed design required, but also re-applications of the techniques after even the smallest design and/or task changes are made. Typical output data are error probabilities to be integrated with the system model. The methods follow broadly similar steps: (1) Analysing the task (i.e., what is the human supposed to do?), (2) Identifying potential errors (i.e., where can this go wrong?); (3) Selecting the most significant errors (i.e., which ones are critical to system safety?); (4) Assigning probabilities to the human errors; and (5) Integrating the results into the system model to assess overall system dependability.

The major areas of concern are: (1) The methods cannot explain why errors occur because they focus on the external appearance of errors and do not address their underlying causes; no mention is made of the psychological causes of errors (*e.g.*, decision making is hard to represent in a task analysis, thus several human reliability analysis techniques ignore this cognitive activity entirely). Where it is considered, it is usually treated as a separate analysis; and (2) The methods cannot predict system breakdown. Generally, error identification at the sub-task level has not helped to predict system breakdown (*e.g.*, probabilistic risk assessment techniques assume independence between system events and may, therefore, miss pathways toward failure.)

Performance shaping (or influencing) factors are an important addition to human reliability analysis techniques. After a task analysis has been conducted and basic error probabilities are assigned to the various tasks, these may allow a designer to alter the probabilities in a meaningful and repeatable way. Many factors are known to affect human performance; among the most notable performance-influencing factors are: (1) time pressure; (2) information quality; (3) procedural quality; (4) task complexity, and (5) operator training. In human reliability analyses, such factors are included as independent variables. Little guidance exists as to what factors should be taken into account and precisely how much they would affect the probability of an error.

3.10.2 Human error is dependent upon processing mode. The kinds of errors a human might make in executing a task depend upon the interaction mode. Errors are identified by tracing the three different modes of interaction proposed by Rasmussen (e.g., skill-based, rule-based and knowledge-based. The human operator is no longer seen as a passive element in the system, and errors are classified on the basis of underlying psychological processes. Typical input data are tasks classified by mode of interaction. Output data are the forms of errors associated with tasks that involve different modes of interaction. Error prediction methods that are based on processing mode are limited by the fact that they cannot deal with different levels of operator experience. (Parenthetically, neither do task-based methods.) In classifying tasks according to the mode of processing or interaction, the designer must assume the competence of an *average* operator, since the cognitive mode in which a task is performed depends on both the task and the individual performing the task, and may require input from subject-matter experts. Other problems include: (1) Different levels of processing may run in parallel; (2) It is unclear how finely a task should be broken down before task classification by skill, rule, or knowledge level is performed. and (3) Classification by processing mode will not yield numeric error probabilities.

3.10.3 Human error is the product of a mismatch between problem-solving demands and resources. Although the inclusion of cognitive resources allows designers to trace and predict different kinds of errors according to processing mode, it is less clear why such errors might occur. The demand-resource mismatch perspective takes the cognitive approach further and seeks to explain the reasons for problems in human-machine interaction. Typical inputs are from practitioner knowledge about task demands. The output data are pointers to areas where problem demands outnumber resources. The method involves two steps: (1) Identifying the demands placed on the human in a problem-solving situation, including the knowledge necessary to generate the right problem-solving strategies, the attention that must be distributed efficiently across the operational world, and the goal conflicts (*e.g.*, safety *vs.* production) that must be resolved on-line; and (2) Identifying the degree to which the human-machine system provides the resources to meet these demands. Limitations of this approach include the fact that the methods cannot be driven by enumerations of actions according to event-tree or processing mode. Instead, identification of potential human performance problems is fundamentally problem-driven. Sequences of human actions and system responses (and vice versa) are examined for their potential for interaction problems. However, domain experts may be unable to provide

designers with exhaustive enumerations of difficult domain problem scenarios.

3.11 Micro Models

Micro Models, often based upon a large body of empirical data, have been developed for many different performance variables. For example:
Single finger keying = 0.140* number of keystrokes
Choice reaction time = K*Log(n+1) where K = 0.4983 and n= no. of alternatives.

MIDAS and IPME incorporate a large number of such micro models, in an attempt to model overall system effectiveness as discussed in para 3.12.

3.12 Integrated Models

Integrated HPMs typically attempt to address the human, the physical system and the environment, and by their nature, such models are internally complex. Thus, their validity may rest heavily on the way in which the components interact. Although few such models have been developed, the most notable examples are MIDAS and IPME.
Integrated models have the obvious advantage of treating human performance holistically. Their potential drawback is that, if the model is deep as well as broad, significant effort may be required to use it for even relatively trivial applications. Verification and validation are problematic for integrated models, because of the large number of interacting parameters. A useful distinction may be made between physical and functional integration. The former uses the output of one model as the input of another; the latter is based upon a common underlying cognitive architecture (such as attentional processes represented as an undifferentiated resource pool) that determines the requirements for individual models. The component models do not need to be at the same level of granularity, however. For example, a detailed model of vision may be used with a simple model of overall operator workload. Alternatively, the Model Human Processor (MHP) described by Card, Moran and Newell comprises perceptual, cognitive, and motor systems, each of which contains memories (with an associated capacity, decay and type of code) and processors (with individually specified cycle times). The psychological literature was used to provide estimates of these parameters and established micro-models, such as Fitts's Law that relates movement time to target size and distance, were incorporated. Using MHP, an operator task can be decomposed into its component parts and an estimate derived of the overall level of performance.

3.13 Models in HOMER

Table 1 lists and categorises the models contained in the prototype version of the expert system described in Section 6, according to the factors described above.

MODEL	TYPE	INPUT	PROCESSES	OUTPUT
OCM/ PROCRU	Control	Task dynamic, noise parameters	Kalman Filter predictor, neuromuscular	Real time continuous control
ORACLE	Sensory	Task, environment & observer characteristics	Perceptual rules	Absolute target acquisition performance
JACK	Anthro.	CAD files (workspace dimensions), Human anthropometric data	Force, vision envelopes, limb mobility envelopes	Vision and reach envelopes, collision points
TAWL/ TOSS	Workload	Tasks, times, loads (VACP), subsystem used, operator interdependencies	Overload, workload summary	Workload measures
Win-CREW	Workload	Tasks, sequencing, decision logic, interdependencies	Micro-models, workspace layout, task loadings, operator strategies	Time/error, operator status
W/index	Workload	Tasks, crew station configuration	Attentional limits	Attentional demands
PHRASE-2	Human Error	Human and Machine checklist	Error database, error calculations	Error rate
MIDAS	Task Network	Graphics Files (Cockpit world), task/subsystem list,	Cognitive models, vision models, Jack,	Dynamic visualisation of sys performance task, timeline,

		co-ord (Jack, workstation)	scheduler	workload values, reach envelopes, visual field
PUMA	Task Network	Task Loadings, Scenario	Workload algorithms, library of scenarios	Workload vs. time task timelines
IPME	Task Network	Single task ratings, task networks	Combining rules	Dual -task, performance and workload
HOS	Task Network	Tasks, sequencing, decision logic, interdependencies	Micro-models, workspace layout	Time/error, operator status
FAIT	Task Network	Human/Machine environment	Information Flow model	HF issues re questions, trade-offs and scenarios

Table 1. Names and characteristics of HPMs included in the prototype version of HOMER.

4. MODEL LIMITATIONS

While HPM technology has advanced significantly over the past twenty years, there are still areas in which HPMs have limited capability. This chapter identifies known limitations of current models and is included to suggest areas where model development would have high potential payoffs.

4.1 Co-ordination with Other Sources and Scales of Performance Simulation

There is a potential mismatch between representations of human performance provided by models of individuals or small groups and large-scale simulations of many human and system elements. The main source of the difference is in the size of performance prediction of interest and differences in measurement between large-scale, individual and micro-behavioural models. This mismatch takes several forms, as described below.

4.1.1 One mismatch exists between the level of prediction offered by models of individual or small team performance and those of large-scale integrated system performance; the predicted output of the individual and the interaction among individuals in small groups have a common frame of reference in terms of world information and the information dynamics of that shared world. At some point, when the group of individuals becomes sufficiently large (and it would be interesting to understand analytically the point at which this occurs and its dimensions), the rate and density of information that needs to be communicated and the level at which performance can be predicted shifts. Identifying information bottlenecks in distributed command and control networks is not well understood and representing the dynamics of large-group information flow is beyond the scope of current human modelling.

Another mismatch is the level of performance representation and prediction between the individual model and the large scale system model. The difficulty is that it is not always possible to aggregate the contribution of individual performance to overall system effectiveness. This is especially true for team performance, in which the contribution of the team to success or failure cannot be easily attributed to its constituents.

There is also a mismatch between the level of data provided by micro-behavioural models (either performance or neurologically based models) and the observable performance of operators in either real world or simulated operation. Hence, a model of selective visual attention that predicts a stimulus onset asynchrony of 40 and 50 msec does not generalise well to the level of performance of visual search.

Status:
The granularity and scalability problems are likely due to a lack of sound research, (*i.e.,* a lack of knowledge and data) as opposed to a limit in the state of the art of modelling architecture. Both the development of massively parallel computing platforms and the development of high level architecture should support the representation of multiple interactive agents at whatever scale is desired. The behaviours of interest and the critical performance phenomenon are unknown, however.

4.2 Predictive Decision Making Models (of both Individual and Team/Distributed Decision Making)

Predicting the course of action a human will follow during a moderately complex task has proved to be a difficult modelling task. That is to say, the development of accurate and reliable *predictive* models of human cognition and decision making has proved to be very difficult. Optimal task selection algorithms do not predict, typically, the decisions that humans will make. Rather, heuristic models of human decision making have proved to be useful as explanatory tools. First, these models are expressed as computational algorithms only rarely. Additionally, the analyst must guess which heuristic an operator might use in a given context and then make an appropriate assignment of weightings to those heuristic combinations of factors. This prediction about the *process* of decision making makes the accurate prediction about the *outcome* of decision making problematic. Other more descriptive process models (e.g., recognition-primed decision making) do not, as yet, have the computational rigour to be integrated into human performance predictions.

There is a set of decision-making models at sensory and perceptual levels that are structured as parallel distributed computational network representations. These "neural-net" decision models do successfully predict perceptual decisions if given a sufficiently large and generalised training set. They are, however, not amenable to the explanation of that decision behaviour in terms of reference beyond the model formalism (e.g., node weights, propagation structures.).

Status:
It is not believed that this lack of predictive decision models (especially in a constrained domain with "optimal" operators) is a fundamental limit in human performance representation. A rigorous computational model framework and a set of "situated" empirical studies would likely contribute a great deal to our knowledge and ability to represent decision-making behaviours.

4.3 Representation of Affective or Motivational States

Level of motivation, confidence in performance and "leadership" variables are known to be critically important in most stressful environments . While there has been a considerable attention devoted to teaching appropriate motivational strategies and "crew resource management," there have been limited attempts to integrate affective and motivational state data into a computational representation of human performance. The kinds of effects that might be predicted as a function of motivational state are likely to be of the form accounted for by performance shaping factor structures (i.e., broad changes in a consistent direction across a wide range of tasks). However, the computational framework to describe these changes is not available.

Status:
To reflect human performance adequately, especially at the extremes of behaviour, some method of incorporating motivational and other affective mechanisms is required. This is especially true in the development of models of team or small-unit interactive behaviour. The gulf between research into social and interpersonal behaviours and the computational frameworks that have been developed during the same period of time is extremely large. A small effort undertaken by the National Aeronautics and Space Administration in the US to account for and model "motivated cognition" will be explored.

4.4 Learning as an Active Model Component and Training as a Measurable Modelled Procedure

One of the issues in representing training in a computational model-based simulation is that the temporal horizon for the simulation is measured in minutes to hours while training and learning occur over days, months and years. That fact not withstanding, a recent analysis of distributed interactive simulation for wargaming has stated that the lack of adaptability and learning in the SAFOR and OPFOR representations was a critical shortfall in the acceptability and face-validity of their operation . There is a fairly extensive database on the effects of practice on learning procedures and developing automaticity of operations. There is also some research into training effectiveness to specify training requirements and proficiency levels. It is believed there is a sufficient body of knowledge to support the development and implementation of the consequences of training on performance in a computational form. However, there has not been, to our knowledge, a focused effort to combine the principles and data that are available to compose a predictive model of human learning and training impact.

In another approach to the same issue, there has been a considerable amount of research into the development of effective, computer-based training systems and modelling the "learner" in adaptive training systems. Though there is no evidence for fundamental breakthroughs in this area, it is worth exploring as a source of ideas for human predictive behaviour in training situations.

Status:
A body of data and a system of evaluation for training systems that might be sufficient to support computational modelling of learning and training effectiveness in tightly defined performance ranges appear to exist. Further, there may be data to support a computational measure of the effectiveness of training systems in a simulation environment

4.5 Human Scheduling and Procedure Management

Task prioritisation scheduling and procedure management have not been the focus of research in psychological terms that are consistent with HPMs. Most scheduling algorithms have been developed by industrial engineers for creating optimal manufacturing schedules. Memory for behaviour in a dynamic environment, a key factor in human scheduling, has not been studied until recently. The present work concentrates on individual differences and the interaction between the environment and the scheduling process. While this is useful for *explaining* scheduling behaviour, it does not provide much leverage in the pursuit of *predicting* scheduling behaviour. Early work by Tulga and Sheridan (1972) pointed to some of the issues. Scheduling behaviour is critically dependent on the level of expertise of the operator performing the scheduling process. It is very sensitive to pay-offs, perceived risks, and context variables; the schedule and the process are updated dynamically and not only in response to local constraints, but also in response to perceived global success or failure status.

Status:
Because of the heavy dependence of HPMs on predicting the time required to perform an activity (see next item), the lack of a robust and validated human task scheduling mechanism is critical. If the output of an HPM is a time-line, and if the management of that time line is a measure of critical performance, then a lack of reasonable schedule and priority models is a fundamental and significant flaw in HPM.

4.6 Predicting Performance Level and Accuracy as Opposed to Just Performance Time

HPMs have provided, in both the network- and psychological-model-based forms, performance

predictions in terms of time to perform, percent completed performance, delayed performance, and relative performance times in support of comparative types of analyses. However, there has been little or no development of predictive measures that reflect the *quality* of performance on either a given task or the trade-off between performance quality, schedule and load level. Like schedule management, performance management is a hallmark of skilled operation. To be able to predict neither these types of behaviour nor performance quality in either relative or absolute terms, is a critical shortcoming. Some inference about performance level and accuracy can be made looking at the temporal characteristics of behaviour. (*e.g.*, at the time limit, either the task could or could not be performed). However, the explanatory power of such an inference is very low and it misses the relationships between the quality of performance on one set of tasks and the performance on other behaviours.

Status:
There is hope that, as internal representations of operator processes are developed, more diagnostic performance measures can be developed for HPMs. However, the assertion that a model process predicts an internal process has formidable validation issues, unless the predicted behaviour can distinguish between one internal process and another unambiguously. Very few psychological constructs have had success in this kind of differentiation. On the other hand, there have been some practical computational approaches developed for modelling the interactions among tasks given performance times and accuracy rates. However, their theoretical underpinning has not been established yet.

4.7 Predicting the Variability of Human Performance in Addition to Mean Performance

There are many features that characterise the performance of a task. Even constraining our measures to fundamentally temporal ones, the predictive representation can be improved by manipulating the characteristics of the performance distribution. Variance curve type, scatter, kurtosis and cut-offs can all be successfully manipulated to produce accurate variations in the human/team's performance. In addition to these degrees of freedom, there is likely a fair amount of data available to characterise the appropriate variations.

In studying human performance, it is often the variability that is of great interest. Human performance is highly variable relative to most other system design elements. Therefore, the designer should consider performance variability as well as average performance.

Status:

It is believed that the information on performance variance is available and can be easily adapted to serve both network and principle-based models of human performance. This is one dimension on which progress can be made immediately.

4.8 Situated Cognition Affordance and Ecologically Valid Situation-Sensitive Performance Models

Few HPMs have a well-articulated representation of the environment and equipment with which the operators interact and fewer still have included that representation as a driver for behaviour. The integration of both constraints and performance leverage in the interaction between the operator and his operating environment is a critical part of human performance modelling. About 10-12 years of research have been performed in the area of "ecologically valid" performance and situated cognition. The methods that have been brought to bear in this research does not yield the structures and performance variables that have been used to guide HPMs. However, there may be a sufficient data set at this point to begin to articulate the impact of "general affordance" on behaviour. Again the type of work performed has tended to support more of the *explanation* of behaviour than the *prediction* of behaviour, but regularities may exist that can be exploited. These situated decision and cognitive performance models may also yield performance measure and performance shaping factors that have not been previously exploited.

Status:
There have been a couple of efforts to describe characteristics of situations either as sets of states of the environment (e.g., phase of flight) or by describing the operational procedural chain. Inclusion of more of the factors of environment and equipment in these kinds of descriptors may move the HPM in the direction of more "ecologically valid" performance measures.

4.9 Assessing the "Coverage" of a HPM

It is a fundamental truism of modelling, regardless of domain or focus, that 'all models are wrong, but some models are useful.' All models are wrong because a model is not reality - - it does not fully represent the reality that it models and thus, it will be necessarily an inaccurate representation of that reality. Nevertheless, some models are useful because they have included important and relevant parameters in a package that is less complex (and, therefore, more manageable) than reality. The problem, inevitably, is in keeping track of

the "coverage" of the model i.e. which of the important parameters have been included in the model and which have been left out. It is important to ensure that final decisions are made on the basis of models and analyses covering all of the important parameters for the problem.

This is particularly problematic for task- or scenario-based HPM tools (which comprise the majority at the current time). Since it is impossible to model all but the smallest subset of the scenarios in which a system will be used, it is important to ensure that the set of scenarios that are modelled cover the space of possibilities. Even so, a fundamental critique of this approach to system development, is that it will be the unexpected scenarios that will prove to be disastrous, as they have proven to be in accident after accident in the past.

This implies that there should be some sort of overall understanding of the problem space against which the HPM user can ascertain the degree of coverage provided by a model. This "problem space model" would seem to have at least two important dimensions: environmental factors and behavioural factors. Good coverage of environmental factors means that all aspects of the context which can affect human-system performance have been covered. This could include everything from visibility conditions and sun spots to system failure modes to human mental models about the environment and even human physical characteristics. Good coverage of behavioural factors means that all aspects of those actions which are possible in the environment and which can affect human-machine system performance have been covered. This becomes the set of action-based scenarios (including the actions of the human, machine and external world actors) which have been examined. For both dimensions, it should be noted that understanding what portions of the space have been not been examined may be nearly as useful as ensuring good coverage.

Status

Accurately and reliably assessing model coverage is a fundamentally difficult problem, especially for novel systems, because it requires a more complete understanding of the domain than usually exists. Most progress on this front has been made by providing "reference models" against which coverage of the HPM analysis can be assessed. For example, the FAIT technique uses a reference model for the classes of interaction between human controllers, machines and automation, and the environment (called the 'mixed initiative model') to provide a conceptual check on behavioural coverage., Thus providing a measure of assurance that considerations at all points in the behavioural interaction cycle have been examined. Similarly, Rassmussen's Abstraction Hierarchy has been used to provide a conceptual check on environmental coverage, providing a guarantee that all potentially important aspects of the environment have been included in an analysis, to at least at some level of detail and granularity. More work needs to be done to understand these problem space models and the activity of modelling needs to be more closely integrated with these reference models in a fashion similar to that used for requirements tracing in software development currently.

5. IMPLEMENTATION ISSUES

This Chapter addresses some of the pragmatic issues associated with the fielding of HPMs. It is intended as guidance for HPM developers to help ensure that their models will be usable by systems engineers.

5.1 Defining the Scope of a HPM

Central to the use of any model is the issue of scope and limitations; models fully applicable to one type of problems may be entirely inappropriate or inefficient for the other. For example, Newton's second law, force = mass x acceleration, is only appropriate to bodies travelling substantially slower than the speed of light. For higher speeds, a different model would be needed that could account for the effects of relativity. Similarly, a model designed for predicting human workload may not be appropriate for predicting the consequences of decision making strategies in command and control.

In assessing the 'goodness of fit' of an HPM to a specific design issue, it is implied that criteria or objectives exist to answer the question "Goodness of fit to what?". These are the criteria that need to be established before labelling a model as usable or not.

5.2 Integrating Human Performance Modelling into the Systems Engineering Process

While human performance is often a high-risk element in overall operational effectiveness, the traditional design process tends to focus on the performance of hardware and software with little attention to the human component. Part of the reason for this is the historical lack of HPMs. Now that models and tools are available for inclusion in the systems engineering process, some cultural changes may be required, that might include:

5.2.1 Develop a good understanding of user tasks and goals early in the design process. This should include an understanding of what users will accomplish with the system, the types of tasks they will perform, and the decisions they will make, measures of human effectiveness, environmental conditions, the information required, etc. These data should be used along with that focused on the functionality of other system elements to drive the design process and to ensure that proposed sub-systems are assessed as an integrated whole in terms of their ability to work together to support user tasks. This viewpoint provides the foundations from which system models, including operational analysis models, are built and from which equipment and HPMs are derived.

5.2.2 *Develop/employ metrics for evaluating human performance as a component of system performance.* These metrics are essential in the us the data generated in the models to influence the design. The scenarios in which the system will be used, and the human performance which should be achieved should be defined to the extent possible. These metrics can be established at a top level early in a project, and elaborated as design detail emerges.

5.2.3 Place greater emphasis on human error. More attention should be paid to conducting human reliability analyses as part of the system risk analysis. Human error should be included in failure modes analyses and safety hazard analyses to ensure that an error tolerant system is developed.

5.2.4 Use models to identify where human performance is critical to mission success. High fidelity models of human performance are expensive to build. Therefore, it is important to be selective in identifying areas of high risk so that the modelling and data collection resources are best allocated. Lower fidelity HPMs in conjunction with models of other system components provide the tools to focus these analyses.

5.2.5 Generate/collect human performance data to "feed" model for areas of high risk. Significant resources are often spent producing data to refine models of equipment performance. In a similar manner, human performance data may be needed to improve HPMs. The cost-benefits associated with collecting and analysing human performance data should be considered during project planning. Furthermore, mechanisms for reusing these data between projects should be developed. Companies may realise returns on investment associated with building up libraries of human performance data for use in HPMs to support design trade-offs.

5.2.6 Make use of prototypes and simulations standard practice in system design. Prototypes and simulations involving human operators give users a chance to "test drive" the system. In addition, both users and designers get an early view of an integrated system, which can greatly enhance system usability. However, the role of simulation can and should be extended to help provide objective criteria for the fitness for purpose of systems. Also, HPMs should be used to extrapolate from human-in-the loop simulations to examine human performance outside the narrow conditions of the simulated environment.

5.2.7 Develop improved definitions of the human-machine interface.

In the past, design specifications have focused on the technical performance of system equipment. Now, a definition of the displays and controls with which the user will interact is essential. This allows for a more accurate assessment of the likelihood of human error and/or the time required to perform tasks. In addition, clear definition of the human interface provides a valuable tool for soliciting feedback from users about the emerging system design.

5.2.8 Include human performance in system test. More and more, customer's are mandating the provision of evidence to demonstrate that human factors have been considered in the design process. The output of HPMs can provide this evidence. Furthermore, they can facilitate the development of human performance acceptance criteria to be used in system test.

5.3 Validation of HPMs

Increasingly, HPMs and modelling tools will need to be subjected to model verification and validation (V&V) scrutiny. Particularly in the military domain, formal V&V is essential if model results will be considered in the decision-making process.

Generally, model V&V involves the *verification* phase where the question is whether the modelling software behaves as it is claimed to behave (e.g., algorithms are implemented correctly, random number generators produce truly random numbers). The *validation* phase focuses on the ability of the model to provide sound predictions. Central to validation is defining the scope of the issues that the model can and can not address.

V&V of HPMs poses some unique problems in comparison to that of other types of models. First, and most important, is the high degree of variability in the behaviour of human operators. Unlike hardware and software, the range of performance found among qualified human operators can differ by as much as 100%. A range of 20-40% is typical. Therefore, a large sample of empirical human performance data is required to get a stable estimate that can be compared to the model. Additionally, human performance data tend to be difficult and expensive to collect. Collectively, this means that traditional predictive validation studies for validating HPMs will be rare.

To validate HPMs, it is recommended that other types of validation be pursued in addition to predictive validation: (1) *Face validation* — do the modelling strategies look reasonable and appropriate to the kind of analysis? (2) *Construct validation* — have some of the components of the model (e.g., the workload prediction component) been proven valid through empirical research?, and (3) *Concurrent validation* — does the model predict performance of known and previously studied human systems?

Finally, the ultimate measure by which any model's utility is evaluated is the *value added* to the analysis by that model. As with other engineering and systems prediction models, if they add value to the analysis, it can be claimed that they are worth using.

5.4 Commercialisation of human performance modelling Software

Human performance modelling software is often developed by groups of specialist engineers or scientists, working within larger programmes funded by governmental or quasi-governmental agencies. In these cases, the software may be created to support R&D activities in the first instance, and only later considered for wider release.

Software developed for R&D purposes is quite different from that created for commercial purposes. Consequently, for the human performance modelling software developed in the R&D environment to become commercially viable, issues of software maintenance and support must be addressed.

The term "maintenance", as applied to software, actually has several meanings: (1) The correction of errors ("bugs") in the software; (2) Enhancements to the software to extend its functionality; and (3) Modifications to the software to enable it to run with the latest hardware, or new versions of a computer's operating system.

The manufacturer of a commercially available software package will normally dedicate resources to the above, and from time to time issue upgraded versions of the software, incorporating all of the solutions for "bugs" found since the last release, and any functional enhancements that have been added. Normally, such software releases will be provided in the context of a maintenance agreement, perhaps free for the first year and renewable annually thereafter for a fee amounting to 10-20% of the purchase price of the software. Major upgrades will not be covered by this fee, typically, but existing users will get a discount on the new package. In some cases, where a particularly critical bug has been discovered, the manufacturer may be prepared to issue an interim "patch" to allow the software to run properly, pending the next formal release.

The term "support," as applied to software, typically refers to the provision of a service providing advice and help to the user. Such support may be available via a telephone help-line, or by fax or e-mail, with a guaranteed response time measured in hours or days.

All of the above have proved to be not merely desirable but essential if software is to remain in serious use over a period of time by anyone other than the group that originated it. Accordingly, it is recommended that those who support the development of HPM software give consideration to commercialisation issues when prioritising an R&D programme.

Normally, it will be in the interests of any organisation that creates software for the commercial market to engineer it so that it is well structured, documented, and engineered, since doing so makes it easier to maintain. In some instances, software with an R&D origin may be crafted with other priorities uppermost in the minds of the developers, such as getting the software constructed rapidly or achieving a high level of performance. They may also hold the view that the software will be short-lived, and modified only by themselves. Thus, design and documentation issues may be given, quite legitimately, low priority. The result may be, however, code that is harder to maintain. This working group can only urge that those involved in the software creation process give thought to the notion that some software lives on for much longer that originally envisaged, and so attention to its structure and documentation may be of benefit to others in the future.

It is also recommended that governmental organisations involved in the creation of specialist software recognise the value of commercialisation and actively support the process whereby specialist software is made available to the scientific community through these organisations for the benefit of all. In cases where this route is not appropriate, yet the software is of value for research purposes, it is recommended that the originating organisation makes the source code freely available to users., This could be achieved via the Internet, in the anticipation that maintenance is undertaken on a self-help basis by whatever community of users evolves.

5.5 Model Tool Usability

As with any modern software intended for a wide base of potential users, software usability from a software design perspective should be addressed seriously. Many of the current tools are cumbersome and unnecessarily complex and could be improved through the use of software usability design practices that are common throughout the commercial software development industry. A reasonably coherent overview of software usability, particularly with respect to life-cycle development and the iterative nature of usability testing, is given in Chapter 3 of "A Guide to Usability" (DTI, 1990). The recommendations summarised below are only as general pointers.

5.5.1 Recommendations concerning the human performance modelling environment

5.5.1.1 Input data

- Ensure that the model does not require input data that may be difficult or impossible to obtain. For example timeline data may be required but not available from the requirements-capture phase of the design cycle.
- Ensure that the data format is clearly identified, with respect to units, precision required etc.
- Use internationally agreed upon units, where relevant.
- If transformations of raw data are needed, indicate how this can be achieved.
- Indicate how the model treats missing data (e.g. if a user has 95 % of the necessary data, including all critical data can the model still be used ?)

5.5.1.2 Output data

- Ensure that the format of the output data is specified, so the user can check compatibility with the end application.
- Where possible, permit options for saving data in various formats to allow maximum portability across software and hardware platforms. For example, data export can be promoted by providing save options to generic text files or common graphics standards files (GIF, TIFF etc.).

5.5.1.3 Documentation

- Context sensitive, on-line help is desirable, but note that extensive help may reflect an admission of poor usability and may point to the need for a re-design.
- Users may find it particularly helpful to have sample input and output data files available to allow walk-throughs —it can be reassuring to use the input data to produce data that match the output sample by running the model.
- Consider carefully the role of an operating manual; it may be best to have separate documents for the overall description of the model and the step-by-step guide.
- The overall model description should include a list of the critical assumptions.

5.5.1.4 Hardware Platforms

The development platform should be as widely available as possible. Although some models require significant computing power, it needs to be acknowledged that there is an increasing dominance of a small number of operating systems (including Windows ™ and Windows NT™) that most users will be most familiar with. This will have implications for the promotion of usability.

5.5.1.5 Effort required to use the model

Specify the time required to use the model. A model may be regarded as unusable if it cannot be run within a certain time period, irrespective of how many resources are devoted to the exercise.

5.5.1.6 Modularity

Where possible, try to ensure a flexible, modular approach to the modelling environment. This facilitates responding to future user requests to change the model.

5.5.1.7 Design for errors/delay

People always make errors, thus the model environment should offer reversible actions (e.g., an "undo" function) and good error messages. If a lengthy calculation or a batch job is in progress, inform the user.

5.5.1.8 Consistency

Try to be consistent in the overall 'style' of the model presentation.

5.5.2 Recommendations concerning the user of HPMs

5.5.2.1 Specify the knowledge required to operate the model

Iterative design, as mentioned above, should proceed hand-in-hand with the involvement of a set of representative users. Modelling user knowledge is extremely difficult. It may be necessary to undertake a Task Analysis, approaching the development of the model like any other piece of software to be developed. User knowledge capture should cover knowledge of the domain to which the model applies, computing knowledge necessary to operate the software, and the environment in which modelling is likely to be undertaken. This will help to determine the suitability of the model for infrequent use — a user without the necessary knowledge is unlikely to be able to undertake modelling exercises without refresher training. At the other extreme, an expert user may wish to have pre-established shortcut keys, or at least a macro facility to permit them to create their own scripts for frequently-performed operations. The expert may also wish to override some of the system assumptions and warnings. Consider also that the necessary knowledge may be available from other sources close to the user.

5.5.2.2 Specify the training needed

If running the model requires a skill that a user may not possess, specify the type and duration of training that will be needed.

5.5.2.3 Provide diagnostic information regarding the source of human performance failures/deficiencies

When a system deficiency related to human performance failures is found, the user always wants to know why, so they can find a way to reduce the likelihood it will occur again. Therefore, the HPM should provide pointers to the underlying cause of the human's failure (e.g., memory overload, inability to monitor two displays simultaneously,).

5.5.2.4 Consider to whom the user has to communicate the results

Typically, the user of models is not the final decision maker on the system design. The model user must often convey the results of the model-based analysis to an engineering design team or managers with less or no formal training in human performance. Therefore, it is important to select terminology carefully and translate analyses into terms meaningful to the rest of the design team and decision-making hierarchy.

6. DESCRIPTION OF HOMER

6.1 Objectives of HOMER

Many models have been developed that have widely differing capabilities, limitations, and requirements along a multitude of dimensions. Thus, it may be difficult for a knowledgeable potential user to consider all of the relevant factors when selecting the HPM most appropriate for a specific application, and almost impossible for a first-time user. To address this need, the working group developed an expert system named the Human Operator Expert Review, or HOMER. The prototype was developed using a commercially available expert system shell made by EXSYS Inc. In its current form, HOMER asks potential HPM users a series of questions about what they wish to do with the model, how much money, time, and other resources they have, what types of output they require from the model and so on. These questions were selected to elicit the types of information that a HPM expert might seek from a potential user before offering advice about the model(s) that might meet his needs. To respond to each of the questions, a user of HOMER is asked to select the option or options that most closely describe his resources and requirements. The options represent capabilities possessed by at least some of the currently available HPMs. Some effort was made to select only those factors that were likely to discriminate among competing models. HOMER then rank-orders the HPMs in its database with respect to how closely each fits the user's requirements, practical constraints, and so on.

The goal was to produce a "living" system that could be updated as new models are developed and the capabilities of existing models are enhanced. The initial version included 13 HPMs that were representative of different classes of models. Each of these models was described and rated by the member of the working group most familiar with the model, in order to develop a proof-of-concept version of the expert system algorithms and philosophy. For later, more complete versions, it is anticipated that the developer of each model will provide the information required to add a new model to HOMER. These responses will be taken at face value and no further evaluation or critique of the *quality* of a given model along a given dimension will be made by those responsible for maintaining and expanding HOMER. Although this limits objectivity, this approach was adopted for practical reasons.

6.2 Description of the expert system shell (EXSYS)

EXSYS Professional is a multi-platform environment for developing expert systems. HOMER was developed with the Macintosh version, although run time versions of the finished expert system are available for both Macintosh and IBM-type personal computers running under the Windows™ operating environment. The rules that comprise the expert system were developed with an If...Then...Else format. Each rule has several parts: (1) a statement that is either true (the user selects it because the statement represents his situation or requirements) or false (the user does not select it). The "If " part of a rule is expressed as a statement (e.g., "My primary interest is..." which the user completes by selecting among one or more variables (e.g., ... crew complement, ... display format & dynamics, ...workspace geometry, etc). (2) in the "THEN " part of the rule, a specific value is *added* to the confidence value for a candidate model (if the model is capable of a function that the user requires) or *subtracted* from it if it is not, and (3) a note that provides the user with additional information about the question at the user's request. EXSYS keeps track of the values each choice receives as the rules are processed and calculates a final confidence value for each choice. Although EXSYS offers forward and backward chaining and the possibility of more complex logic, a simpler approach was adopted for this application. A number of alternative ways of handling uncertain data are available in the development environment; the "increment/decrement" system was selected for this application. Points (whose values were determined by the working group and are reviewed below) are added to or subtracted from the accumulating total for each of the models considered by the system. At the end or each iteration, the confidence values for the top-scoring models are displayed so the user can view those which most closely fit his stated requirements and constraints. If the user wishes to ascertain the impact of changing one or more of his answers, or to review the answers that he gave during the previous run, this can be accomplished easily. For example, a user might be interested in the impact of a larger budget, longer lead time, or less ambitious requirements on recommended models.

An example of the underlying data and structure of the model are shown below:

Rule 1: (IF) My primary interest is crew complement
(THEN)
MIDAS Confidence = 5
PUMA Confidence = 5
.
.

Rule 2: (IF) My primary interest is team interactions (e.g., CRM)
(THEN)
MIDAS Confidence = 15
PUMA Confidence = 15
.
.
.

Rule 83: (IF) The model must generate a dynamic visualisation (animation)
(THEN)
MIDAS Confidence = 16
PUMA Confidence = 16
Phrase-2 Confidence = -16

6.3 Expert system development process

The working group began by generating a lengthy list of questions that an "expert" would typically ask of a naive user. Most of these have been discussed in previous sections of the report. The first and most obvious question to be asked concerns the goal of the analysis or problem the user wishes to solve with the model. The degree to which each model has been optimised for that problem domain is then given considerable weight in computing the final answer. Thus, for example, if a user is most interested in control-system design, the Optimal Control Model would be more likely to satisfy his requirements than would FAIT, other things being equal. Questions about the stage of development and previous availability of the equipment or system to be analysed are relevant because many models require more detailed information about the physical system (e.g., ORACLE) or flow of information and events (e.g., IPME, MIDAS) than do others (e.g., FAIT). Resource questions address practical constraints that may have little to do with either the goals of the analysis or a model's ability to satisfy them. They do, nevertheless, determine whether or nor it will be feasible to procure the software and/or hardware, staff the analysis effort appropriately, and complete the analysis in the time available using a particular HPM. Many questions address the types of input a model will require to perform a specific analysis; one model might require a digitised rendering of a workspace layout whereas another might require a timeline of a typical mission. If such inputs are unavailable, then models that require them are not considered to be good candidates. Similarly, if a particular type of output is required, only those models that are able to provide such information are good candidates. Thus, Jack provides excellent information about reach, fit, and biomechanics for workstation design while offering little information about operator workload or decision making processes. Alternatively, Oracle offers detailed estimates of operator performance with a specific device while performing a specific task, but is inappropriate for analysing multi-crew operations. Many models offer some sort of dynamic output to aid the user in visualising the mission or vehicle under analysis (e.g., MIDAS, IPME) whereas other have no such capability, offering instead various statistics and estimates (e.g., TAWL/TOSS, W/Index). These are only a few of the issues that might be considered in selecting a candidate HPM. The goal of this approach is to ensure that the potential user considers all of the relevant aspects of the decision-making process and is helped to weight them in a meaningful manner. The final set of "questions" became QUALIFIERS in EXSYS parlance.

Next, the working group listed the choices that a user might make given the capabilities and requirements of existing models. These became VALUES in EXSYS parlance. This list was iterated a number of times until the minimum number of questions necessary to discriminate among models was achieved. The dimensions along which a model might be evaluated included: the topics it covers, the types of equipment and stages of design it can represent, practical issues related to cost, hardware and personnel support, the way it handles data and the output it provides. The relevant dimensions are represented by 21 questions or qualifiers in the beta version of HOMER. The number of choices available for each question range from 2 to 15, with the additional option of "not applicable/important" for most questions. In many cases, the user is required to respond with a single choice. However, rerun the model to compare the effect of that change. The questions and values are listed in Table 2.

Depending on the user's response to each question, and other constraints imposed by the working group and represented in the expert system rules, confidence values are assigned to each of the candidate models. The numeric values are based on three factors: the importance of the question (weight), the format or type of question (rating range), and the degree to which a model does or does not possess a particular quality (rating).

6.4 List of models selected for proof of concept version

There were 13 HPMs selected for inclusion in the beta version of HOMER because they represented different classes of available models, such as those discussed in a previous section of the report (e.g., anthropometric, timeline, procedural, etc). The candidate models became CHOICES in EXSYS parlance. Using the increment/decrement method (we can never completely rule out any model nor is it likely that any model will completely satisfy any user so it is all a matter of degree), 83 "rules" were generated, each one of which is a different QUALIFIER/VALUE combination. Table 1 lists the models included in the proof-of-concept version of HOMER.

6.5 Assignment of weights

The working group believed that some questions are more important than others when discriminating amongst models. Thus, an importance weight that ranged from 1 (relevant enough to be included, but not definitive) to 5 (extremely important, definitive) was assigned to each question,as may be seen in Table 2. In some cases, the weight had the effect of enhancing the probability that models that possessed a particular quality would be at the top of the list. In other cases, the weight had the effect of serving as a "show stopper". That is, if a potential user really needed a particular capability, and a model could not support that function, then the model was given such a negative score that positive scores on other factors would be most unlikely to outweigh that one critical failing.

6.6 Assignment of ratings

Developing the philosophy for the rating and weighting schemes consumed a great deal of the working group's time. For example, the group felt that some dimensions were fairly straight forward, e.g., a model can perform a function, output a type of data, or requires certain input. If the user needs a capability, the confidence levels for models that offer that capability are incremented by a specific value while the confidence levels for models that do not are decreased by a similar amount. In other cases, the impact on confidence values is one-directional. For example, the fact that a model costs less than the amount the potential user has available for the modelling effort does not in and of itself make the model more appropriate (hence no positive value is added), although, the confidence value is decreased if it costs more. In other cases, models might possess a particular capability to varying degrees. For these topics, a range of positive values is available, as well as one negative value.

Three types of rating schemes were used, selected so as to be appropriate for specific questions:

Binary1: +4 if a model had the capacity to perform an important function
(used for capability questions only)
-4 if a model could not perform a function or meet a criteria

Binary2: 0 if a model was cheap enough, timely enough, etc to meet the criteria
(used for resource questions only)
-4 if a model could not meet a specific resource criteria

Graded: +4 if the model could perform a function extremely well or produced far more information than the user provided; if it was designed to do that function
+3 if a model could do something or generate information of a particular type well
+2 if a model could do something or generate information adequately
+1 if a model do a function, but with difficulty or can accept input (but simply passes it back to the user with little value added)
-4 a model could neither generate a value nor accept an input, or required more money, time, etc than the user had available

In all cases, however, the impact of these ratings is strongly influenced by the weight that the group assigned to reflect the importance of each question to the overall task of selecting the most appropriate model. For the first set of questions, those related to the goal a user has in considering an HPM in the first place, an extremely negative value is inserted for any model that is not capable of addressing a specific topic. This value, combined with the significant weight assigned to this question makes the user's response to this question particularly crucial. The group felt that it would be instructive for a potential user to run the expert system with one selection, then choose a different option to view the effect this might have on the HPM recommendations. Finally, the group varied the number of alternative responses allowed for each question; in most cases only one alternative can be selected, although, for questions relating to potential model outputs, multiple alternatives are allowed. The range of ratings available for each question and the number of values the user will be allowed to select during any one run are presented in Appendix B1. The working group assumed that assigning appropriate ratings for new HPMs being entered into later versions of the model will be self-explanatory and will not require further fine-tuning of the model. However, iterative testing will continue to ensure the HOMER is providing useful and accurate recommendations. A number of "user" requirements were simulated in order to test the validity of HOMER's recommendations and adjustments to the questions and alternatives were made as required.

6.7 Questionnaire development

A questionnaire was developed to elicit information from the developers of additional models to facilitate their inclusion in future versions of HOMER. It consists of three parts: (1) a brief introduction and background, (2) a request for summary information

about the model to be added to HOMER, and (3) specific information about the capabilities of that model with respect to the 85 question/choice combinations that comprise HOMER's database. Two completed questionnaires are included in appendix B (B2 and B3)

6.8 Plans for the future

The initial version of HOMER was based upon 13 representative models. The capabilities of these models were evaluated by members of the working group who were familiar with the models, but had not necessarily participated in their development or use. Their goal was to provide reasonable "ratings" to further the development of a proof of concept demonstration. The first test of the system was performed by the working group, adopting the perspective of a variety of potential users of an HPM, answering the questions from the perspective of that user, and then evaluating the credibility of the output. Following refinements to the logic, a second version of HOMER was demonstrated to more than 30 experts in the field of HPM. Further refinements were made to address the issues they raised.

In the future, the working group will seek to populate HOMER with information about additional models. Future plans for HOMER include the possibility of mounting it on a Web site to improve its availability. Model developers will be contacted to elicit information about their models, using the questionnaire described above. As with the initial proof-of-concept version, an evaluative approach will be avoided. Rather, items of information about the capabilities of each model in the database provided by the developer will be tabulated and presented to the potential user of the model with an ordered list of the models that are likely to meet his needs. In addition to the recommendations (based upon the self-assessments of the model developers), a brief information sheet provided by the developer will be provided for each model included in HOMER. These will summarise the name of the model, who developed it (or is distributing it), how to contact that organisation, and a brief paragraph describing the key features of the model.

Questions	Choices	# Choices OK	Weight	Range of ratings
My primary interest is...	.. crew complement .. team interactions .. display format and dynamics .. control design and dynamics .. automation .. procedures .. workspace geometry/layout .. communications .. environmental stressors	1	5	+1 to +4 -20
The design phase(s) I will analyse are..	.. operations analysis/research .. conceptual design .. feasibility; dem/val .. system development .. test and evaluation	1	2	-4 to +4
The equipment/ system I will analyse is..	.. off the shelf .. mod of existing system .. a completely new system	1	2	-4 to +4
The crew I plan to analyse is..	.. a single operator .. 2 or more operators	1	3	-4 to +4
Max time available for completing analysis is..	.. days .. weeks .. months	1	4	-4 to +4
The funds available for software purchase are..	.. $0-5000 .. $500-50,000 .. >$50,000	1	4	0 or -4
I am NOT willing to use a..	.. IBM-type PC (with Windows) .. PC or Sun (with UNIX) .. Silicon Graphics .. Macintosh .. any computer	1 or more	4	0 or -4
Available personnel skills include..	.. subject matter experts .. human factors experts	1 or more	2	0 or -4

	.. computer programmers .. modeller/systems analyst			
Available data include..	.. timelines .. task network .. parameters .. analysis of similar system .. model of relevant dynamics	1 or more	3	-4 or +4
The model should represent workload peaks by..	.. mission duration .. errors	1	2	-4 or +4
It is important that the model supports..	.. a vehicle control model .. crew station layout .. state transitions .. system/automation logic .. physical sim of workspace .. view of the external scene	1 or more	4	-4 or +4
The model must run in..	.. real time; scenario based .. faster - (Monte Carlo sims)	1	5	-4 or +4
For decisions, the model must..	.. emulate decision processes & generate decisions .. generate decisions by following user-spec rules .. introduce user-spec decisions at user-spec points	1	3	-4 or +4
For errors, the model must..	.. generate reasonable errors at likely points .. insert user-specified errors at likely points .. insert user-specified errors at user-spec points	1	3	-4 or +4
Model outputs must include..	.. response times .. accuracy estimates .. crew workload estimates .. task list .. task network .. procedure list .. timeline .. function/task allocation .. biomechanical measures .. fit, reach, visual envelopes .. training requirements .. selection requirements .. estimate of sys effectiveness .. maintainability .. data flow analysis	1 or more	5	-4 or +4
The output must be in the form of..	.. real, absolute values .. figures of merit	1	2	-4 or +4
The model must be capable of generating..	.. mission, task, crew summary .. segment-by-segment summary .. second by second events	1	1	-4 or +4
The model must..	.. generate dynamic visualization (animation)	1	4	-4 or +4
The model must estimate the impact on system performance of..	.. human characteristics .. equipment characteristics .. environmental factors .. stressors	1	1	-4 to +4

Table 2 List of questions and Values

7. RECOMMENDATIONS

The report has identified the significant value of having performance models of sufficient validity for evaluation purposes during the early phases of the design life cycle. It is essential to gain an early insight into potential human factors problems and use modelling as a contribution to the overall qualification process. Therefore these recommendations propose key aspects for system designers and users and the creators and distributors of HPMs to consider in order to realise the benefits of a user-centred design approach

7.1 System Engineers and Tool Users

7.1.1 Ensure analyses/models account for the human component which is significant in system effectiveness and life cycle cost.

7.1.2 Develop metrics for system performance from which human performance metrics can be derived and vice versa: ensure that human performance data is in a form that is meaningful to the overall system design process (i.e., perhaps error rates, reaction times, and costs rather than workload or situation awareness metrics).

7.1.3 Develop a detailed concept of use for your system, and use it throughout design to assess fitness for purpose

7.1.4 Use scenarios to evaluate total system performance (human plus integrated sub-systems) — cost-benefit trade-offs among available mixes of humans and technologies. Early (i.e., pre-prototype implementation) use of HPMs may enable consideration of more and/or more radical design alternatives (even alternatives that no one knows how to build yet) — take advantage of this capability if warranted.

7.1.5 Use HPMs to extrapolate from human-in-the-loop simulations to other scenarios, operators, environments etc. Maximise utility of collected human-in-the-loop data by using HPMs to consider what performance might have been like under alternative circumstances (higher fatigue, lower visibility, a less-trained operator, etc.)

7.1.6 Use rapid prototyping and simulation to generate human performance metrics

7.1.7 Develop libraries/databases of human performance data for use on future projects.

7.1.8 Standardise data storage and handling characteristics to allow data exchange between sub-system models

7.2 HPM Tool Creators and Distributors

7.2.1 Either make the system easy to use or provide an appropriate level of documentation, training, and support

7.2.2 Reduce the burden of data collection by offering default values, embedding or referencing potentially useful databases or functions, providing tools that allow re-use of relevant data from one application to another, encourage user groups, archive and distribute user-developed data, models, etc.,

7.2.3 Integrate models with existing systems engineering tools/models

7.2.4 Use standard interfaces to facilitate import and export of data

7.2.5 Validate models wherever possible. Make clear the limitations or range of the validation. Where validation is impossible or impractical, make the lack of validation clear and consider establishing data collection methods to support future validation (i.e., treat the model as a hypothesis and the users of the model as producing data to support or refute the model).

7.2.6 Work towards tools which either provide data in formats relevant to systems engineers or provide translation methods) for transforming human performance metrics (e.g., workload) into system engineering performance metrics (e.g., error rate, performance time).

7.2.7 Any human performance tool to be used outside the lab should obey good software engineering practices: it should be reliable, robust, easy to use, supported with training materials and engineering support, etc. System engineers rarely want to expend the effort to work with laboratory prototypes.

8. REFERENCES

1. A directory of Human Performance models for System Design (1991)
   DRG AC/283 Panel 8 TR/1

2. Human Performance Assessment Methods (1989)
   AGARDograph 308

3. Human Performance Models for Computer-Aided Engineering (1989)

   NRC Elkind, Card, Hockberg, and Messick-Huey

4. Guide for Human Performance Measurement (1991)

   ANSI CENTC 278/WG10 N18

5. Fighter Cockpits for the Future (1993) Adam,EC
   12th DASC IEEE/AIAA

6. Model Human Processor (1986) Card, Moran and Newell in Handbook of perception and Human Performance (Vol 2)

7. A guide to Usability - Usability now (1990) DTI

9. GLOSSARY OF TERMS

Arousal The degree of awareness of the environment

Attention The general, but not highly directed, allocation of sensory-perceptual functions, possibly involving motor functions as well, to a subset of the possible inputs

Anthropometry That field which deals with the physical dimensions, proportions, and composition of the human body, as well as the study of the related variables which affect them

Automation The increased use of mechanisation and or computerisation

Cognition A general term covering higher mental activities involved in the perception, storage, judging, reasoning and output of information

Conceptual design The process of developing the requirements, structure, dimensions, tolerances, and materials to be used for an entity

Control Any device which enables a user to direct the action or operation of some equipment or system

Crewmember A person assigned to perform duty in an aircraft during flight time. Flight crewmember refers to the pilot, co-pilot, navigator, or (where applicable flight engineer

Crew complement The number of operators required to carry out the tasks in support of the operational mission

Data A formalised representation of numbers or characters which have meaning for communication, interpretation, or processing purposes

Decision making The process of evaluating information which results in the selection of a course of action

Dependent variable A variable such as reaction time used to determine the effect of an experimental manipulation

Display design The presentation of data and/or graphics from a system or device in a format designed for human perception through one or more of the senses

Environmental stressor Any condition in the environment which produces stress in an organism, whether climatological, biological, chemical, mechanical, or particulate

Error An inappropriate response by a system, whether of commission, omission, inadequacy, or timing

Error types Categories of inappropriate responses by a system, whether of commission, omission, inadequacy, or timing

Feedback The return of meaningful information within a closed-loop system so that system performance can be appropriately modified

Function allocation The process of deciding how system functions shall be implemented - by human, by equipment, or by both - and assigning them accordingly

Function analysis An analysis of system functions describing broad activities which may be implemented by personnel , and/or hardware and/or software

Goal An objective for which some activity is initiated and sustained

Granularity the degree of precision required when dealing with data sampling

Human characteristics Characteristics of an individual who is involved in the routine control, function, or support of a system or subsystem, but is specifically not involved in any maintenance on that system

Independent variable A variable under experimental control whose effects on dependent variables have to be estimated or controlled

Interface Imaginary surface across which information is transmitted from operator to machine (by controls) and vice versa (by displays)

Maintainability The retaining of a system in, or restoring it to a specified operating condition within a given period of time using prescribed procedures

Man Machine Interface An imaginary surface across which information and energy are exchanged between the human and machine components of a system. The interface is defined by the displays and controls used by the operator/maintainer to control, monitor or otherwise interact with the system

Memory The capacity for mental storage of feelings, sensations, information, movement patterns, and events

Methodology The study of the method, usually taken to mean an integrated set of methods and rules applicable to some goal

Mental workload The amount of mental effort required to perform a task

Mission That designed activity at a particular location which a system is intended to accomplish

Mission analysis A process to determine the operational capabilities of military forces that are required to carry out assigned missions, roles and tasks in the face of the existing and/or postulated threat, with an acceptable degree of risk

Monte Carlo simulation A method used in mathematics, statistics, and operations research to resolve problems by the use of random sampling. The behaviour of a system is simulated by feeding in values of the system variables, and repeating the operation over different sets of values so as to explore the system under a variety of conditions

Normative Pertaining to or establishing of a norm or standard for evaluation

Operator An individual or robot whose functions may include manipulating, supporting, and operational maintenance of a system or piece of equipment

Perception The process of becoming aware of and interpreting external objects, events, and relationships based on experience following the receipt of sensory information

Performance Any result from the measurement of human activity under specified conditions

Performance measure Any objective or subjective instrument developed to evaluate personnel or equipment effectiveness

Procedure Any instruction set or sequence of actions used to accomplish a given task

Procedural development The development of instructions or sequences of actions used to accomplish a given task

Reaction time The elapsed time between presentation of a stimulus and execution of a response

Real time Having essentially no perceptible delay between the occurrence of an event and the knowledge of the event at another location

Reliability The probability that an item will perform its intended function for a specified interval under stated conditions

Scenario Script describing a possible sequence of events and circumstances

Sensory Any system through which information is acquired about the environment

Simulation The process of assuming the appearance and/or behaviour of a real system

Stress The effect of a physiological, psychological, or mental load ('stressor') on a biological organism, which causes fatigue mad tends to degrade performance

System In general a set of items so related or connected as to form a unity or organic whole

System design The process of developing the requirements, structure, dimensions, tolerances, and materials to be used for unity or organic whole

Task A goal-directed composite of related operator or maintainer activities performed for an immediate purpose i.e. in response to a specified input and yielding a specified output

Task allocation The distribution of tasks or task elements between workers and machines

Task analysis A systematic breakdown of a task into its elements, specifically including a detailed task description of both manual and mental activities, task and element duration's, task frequency, task allocation, task complexity, environmental conditions, necessary clothing and equipment, and any other unique factors involved in or required for one or more humans to perform a given task

Task network The network of tasks that represents the activity being modelled. Defines the sequences of task execution, alternate paths through the network, the conditions under which tasks can execute and the effects of task execution on the system

Taxonomy A description of the way in which HPMs can be classified.

Test Carry out a technique or procedure for determining a quantity or performance measure on one or more dimensions for an individual or product

Time line A representation of actions, activities, or tasks in the temporal domain using a horizontal line or bar

Training requirements the total amount of requirements involved in training a new worker or a worker being taught a new task, such as time, curriculum, training media and evaluation means

Validation Demonstration that a test, standard, or other device addresses the attribute that it purports to address

Workload The level of activity or effort required of an operator to meet performance requirements or criteria

Usability The degree to which users can exploit the potential utility of a HPM.

APPENDIX A

This appendix provides some example case studies which describe the process of using the tools for specific system design problems. The intention is to provide a walkthrough of each tool describing the input data required, the process involved in using the data and the resultant output of the tools. The example problem domain and the appropriate models/tools are as follows:

A1	Evaluation of System Effectiveness	IPME
A2	Allocation of Function	PUMA
A3	Anthropometric Assessment	JACK
A4	Human Reliability	PHRASE-2
A5	Automation	FAIT
A6	Target Acquisition	ORACLE
A7	Workload	W/INDEX
A8	Evaluation of System Performance	WINCREW
A9	Automation and Communication Analysis	MIDAS

Worked Example of the use of IPME in the Evaluation of System Effectiveness

Dr. Andy Belyhavin
DERA, Centre for Human Sciences
Farnborough
Hampshire
UK

The Integrated Performance Modelling Environment (IPME) programme was established in 1995 in the UK Ministry of Defence Corporate Research Programme (CRP) under TG5 with the objective of developing a methodology for quantifying the human performance to system effectiveness. The approach adopted to meeting this requirement, was to develop a software framework based on earlier US work, which would permit the description of the human interaction with the system and the environment based on a task analysis approach. The software framework provides the means to simulate the interaction between man and system based on a task network logic flow.

A sample flow is show in Figure 1 for a simple representation of a land based Surface Air Missile (SAM) system.

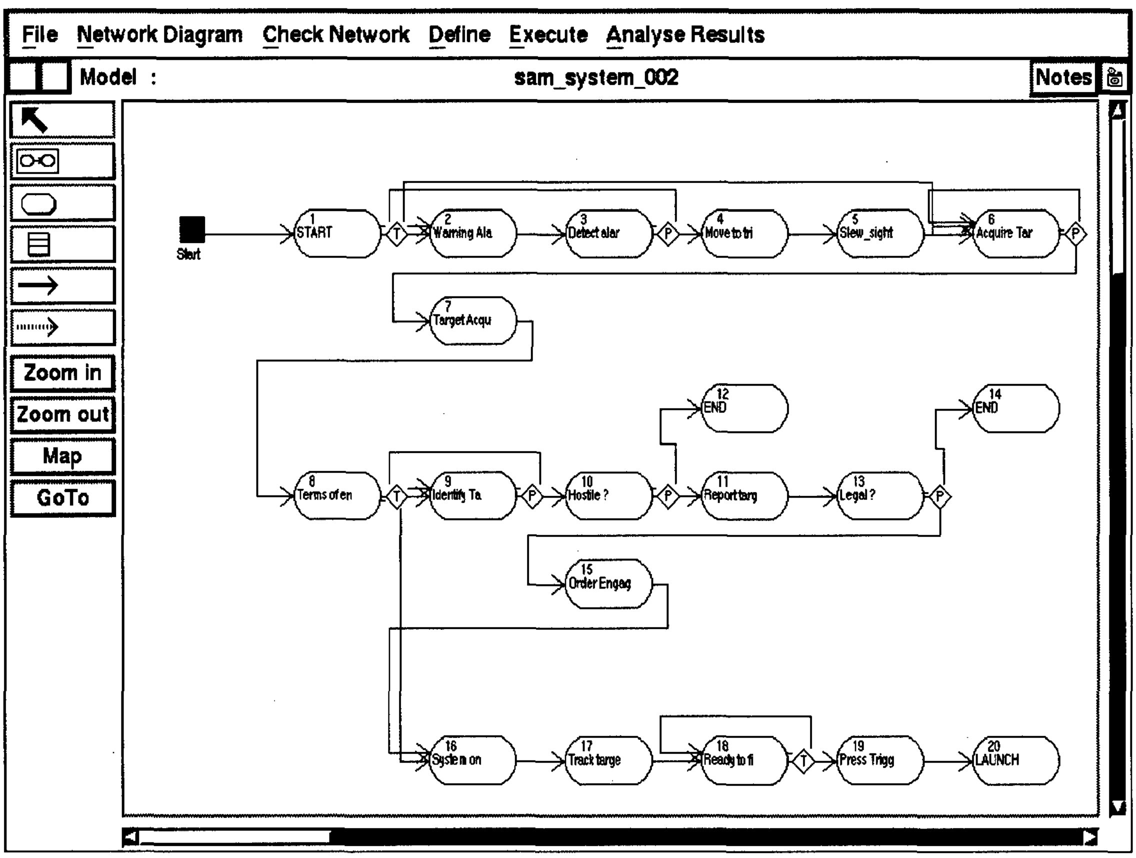


Figure 1: Simple task network flow for a SAM system

The elliptical boxes represent tasks and system activities, the diamonds various types of "decision" box. These "decisions" represent the logical flow of the tasks and can be either human decisions or external events. In the simplified model shown above, the engeagement is broken down into a series of phases: Acquisition, Identification, System on, Target tracking and launch. Within those phases, the processes are represented by loops or parallel activities, which have to be completed before the next phase can start.

The simulation engine driving IPME is a discrete simulation engine based on the US Micro-Saint simulation tool. An event consists of a task starting, a task completing, or the execution of a logic flow decision.

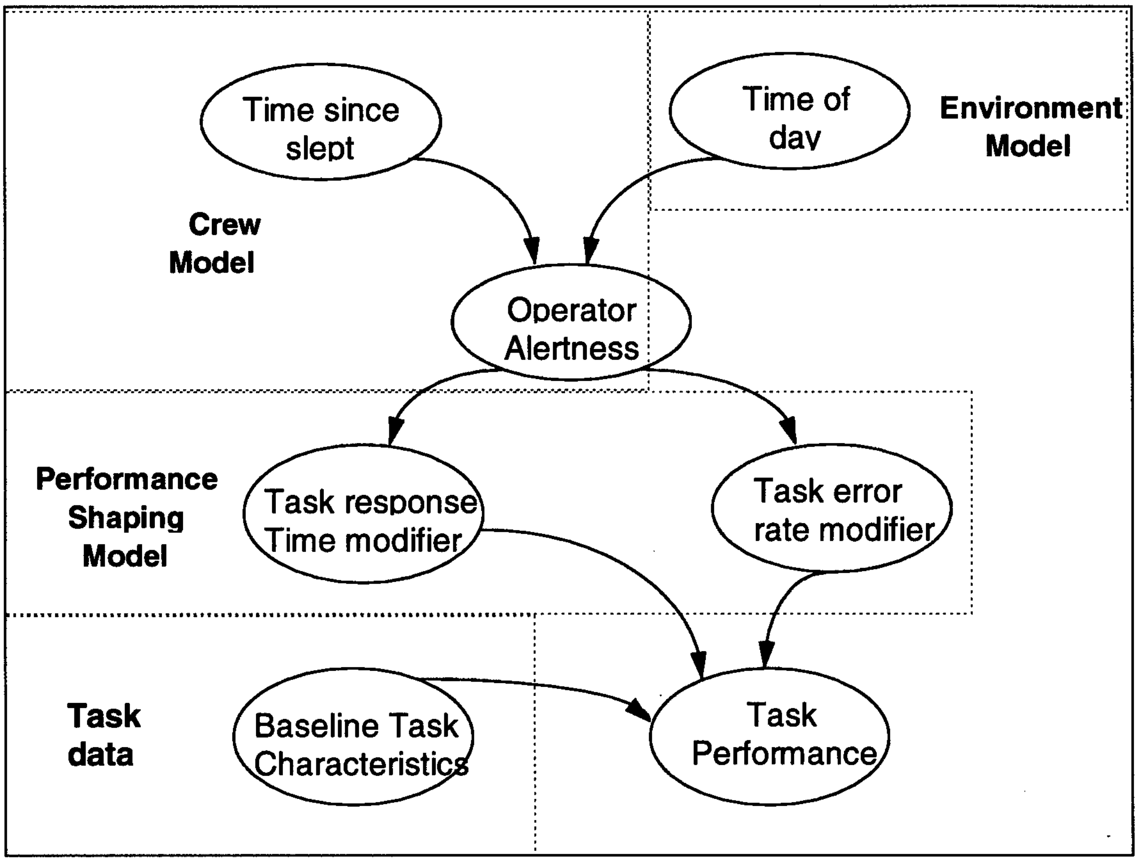
The data required for each task is as follows:

- *Time information.* A probability distribution for the time taken to perform the task. (A task can have a "fixed" time by defining a zero varaince for time to complete)
- *Success information.* The probability that the task will complete successfully (A task can have a zero probability of failure)
- *Failure modes.* The consequences of the task not being completed successfully - e.g. the task is repeated, an alternative task is undertaken etc. (This can be an important component of the system description for hazard analysis)
- *Operator.* Who is doing the task, if an operator is to be involved. "Tasks" can represent both actions of the team and automatic system actions, target movements etc.
- *Nature of the task.* If the task is executed by an operator, it is necessary to allocate the weights in the IPME taxonomy to the task, so that task performance will be modified by the stressors correctly.
- *Task demand information.* If the analysis is to include workload and its consequences, the fields relating to task demand will have to be populated. There are two alternative workload models available in current versions of IPME. The basic version is the DERA Prediction of Operator Performance (POP) model, developed at DERA CHS and DERA AS. The alternative is the Canadian Information Processing (IP) model, developed at the Defence and Civil Institute of Environmental Medicine (DCIEM). Both models require considerable information on task properties, although the data requirements of POP are less than those of the IP model.

To aid the system modeller, there is a library of micro-model times for the completion of a range of low level operator activities, based on well established cognitive and psychomotor theories, first used for the Model Human Processor (MHP) in the middle 1980's and subsequently employed in the US Army Human Operator Simulation (HOS) model.

In addition to the "task" logic flow represented by the network diagram, there is a requirement for background information to populate an IPME system model as follows:

- *Environment information.* This is set up as a distinct model in the IPME framework. It includes models of the behaviour of environmental stressors such as temperature, humidity noise etc., as well as the behaviour of threats and similar external events.
- *Crew characteristics.* These are represented in the Crew model in the IPME framework. A complete team of operators can be represented in the one crew model. The characteristics of each operator are broken down into three groups: Properties (hands / feet / fingers etc.), Traits (height, weight, cognitive ability etc.) and States (TimeSinceSlept, Temperature, etc.). The equations relating these to the environmental variables form a key element of the Crew model.
- *Performance shaping model.* The third of the ancillary models in the IPME framework consists of the functions relating the modification of task performance to the current operator state. It is a basic assumption of the IPME modelling framework that tasks can be allocated to the IPME taxonomic frmaework, and that every task allocated to the same type (taxon pattern) will be degraded in the same way by the environmental stressors or - more probably - through the current Operator state. An influence diagram for the effect of sleep loss and time of day is shown in Figure 2.



Figue B2: Influence diagram for the relationship between environment and task performance for Circadian and Sleep loss effects

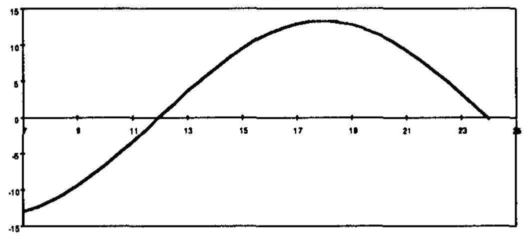
The environmental effects of "stressors", such as duty schedule, are cascaded from the environment and operator models, mediated by the performance shaping model to the final task performance.

In this case, Operator Alertness is treated as a mediating operator state variable. Other environmental stressors which can be treated in a similar fashion are Environmental temperature and humidity, which determine Operator body temperature through Operator clothing, which then determines Operator performance of physical or cognitive tasks. In this latter case, it is not yet clear whether body temperature is the sole determinant of performance, but the principle is similar.

In Figures 3 and 4, the relationship between the Environmental and Operator state measures is displayed for alertness, and in Figure 5, the degradation of successful detections with changing Operator alertness is displayed for a Vigilance task.

Alertness components

CircadianEffects (time of day)

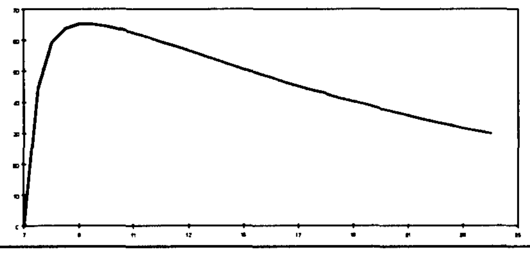


t_{tod} *time of day*

τ *current time*

$$y_{tod} = 13.4\cos(2\pi(t_{tod} - \tau)/24)$$

'S' Effects (time since sleep)



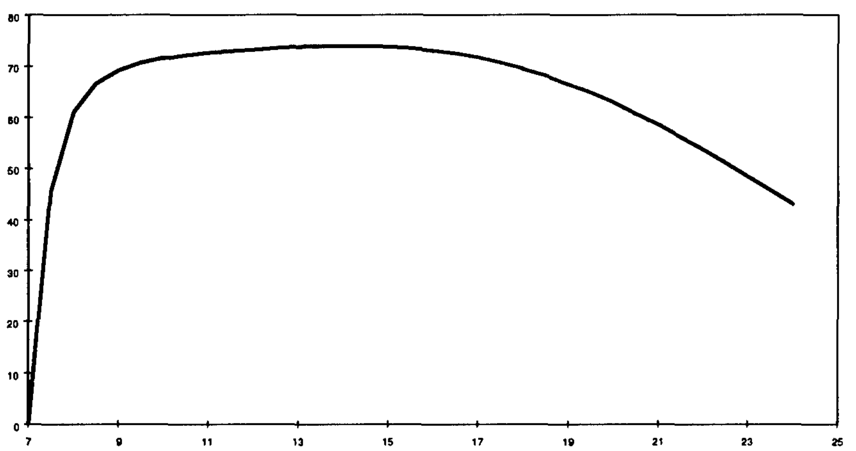
t_{tss} *time since sleep*

$$y_{tss} = 86.5\exp((-0.317/t_{tss}) - 0.0612t_{tss})$$

Source: A model to predict levels of alertness during irregular schedules of work and rest, Montgomery and Spencer 1996 DRA/CHS/A&N/CR/96/007

Figure 3: Variation of Operator Alertness with Time of day and Time since Sleep

Resultant Alertness



$$A = 13.4 + y_{tss} + y_{tod}$$

Source: A model to predict levels of alertness during irregular schedules of work and rest, Montgomery and Spencer 1996 DRA/CHS/A&N/CR/96/007

Figure 4: Resultant Operator Alertness which is the sum of Time since Sleep and Time of day effects

In the IPME framework, the relationship between Time of day and Operator Time since Sleep and Operator Alertness is defined in the Operator model, since Operator Alertness is an Operator state. The final relationship between the state and task performance is defined in the Performance Shaping Model as an appropriate Performance Shaping Function. By way of illustration, the relationship between Alertness and performance on a vigilance task is displayed in Figure 5.

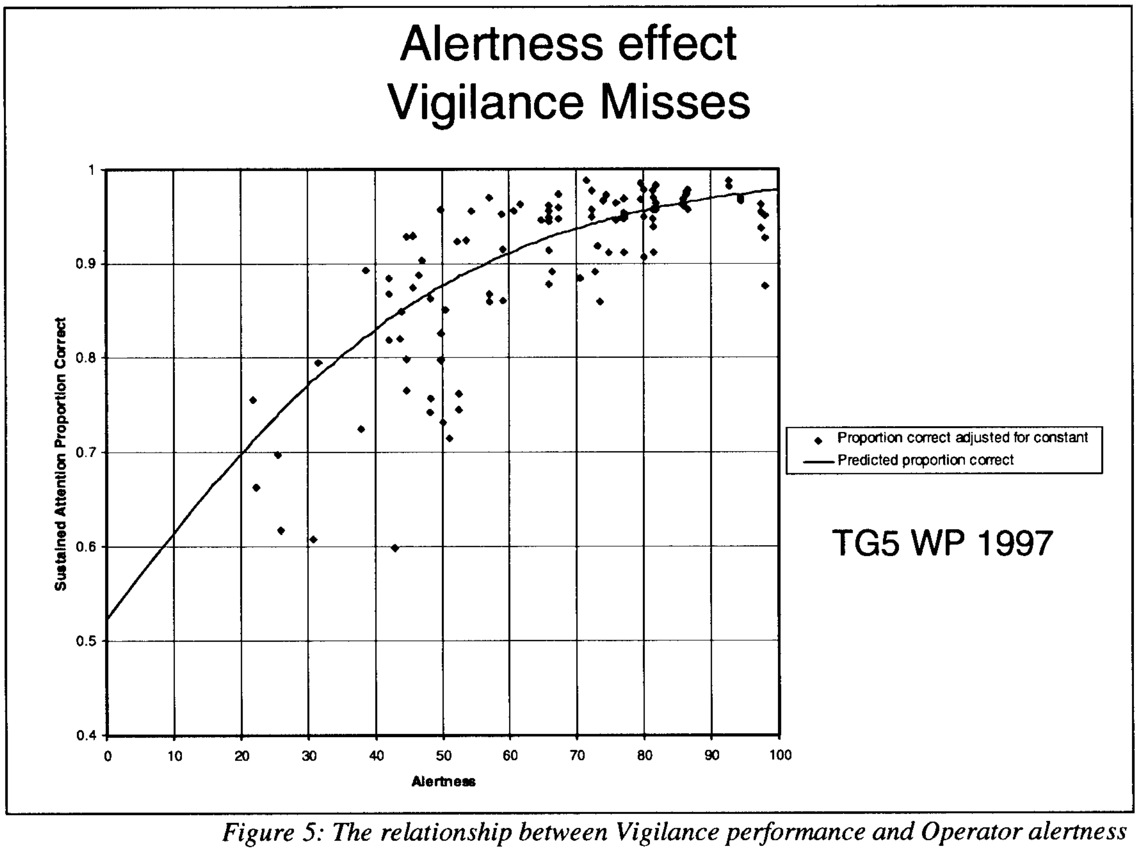


Figure 5: The relationship between Vigilance performance and Operator alertness

In the following sections, a sample of the IPME screens is described for the system displayed in Figure 1.

1. *Opening screen.* The opening screen for IPME provides the access to the 'database' and 'system' screens. In the walk-through, the sequence of screen for opening the Sam_demo system will be demonstrated and the 'Open System' button selected, since a database has been opened automatically.

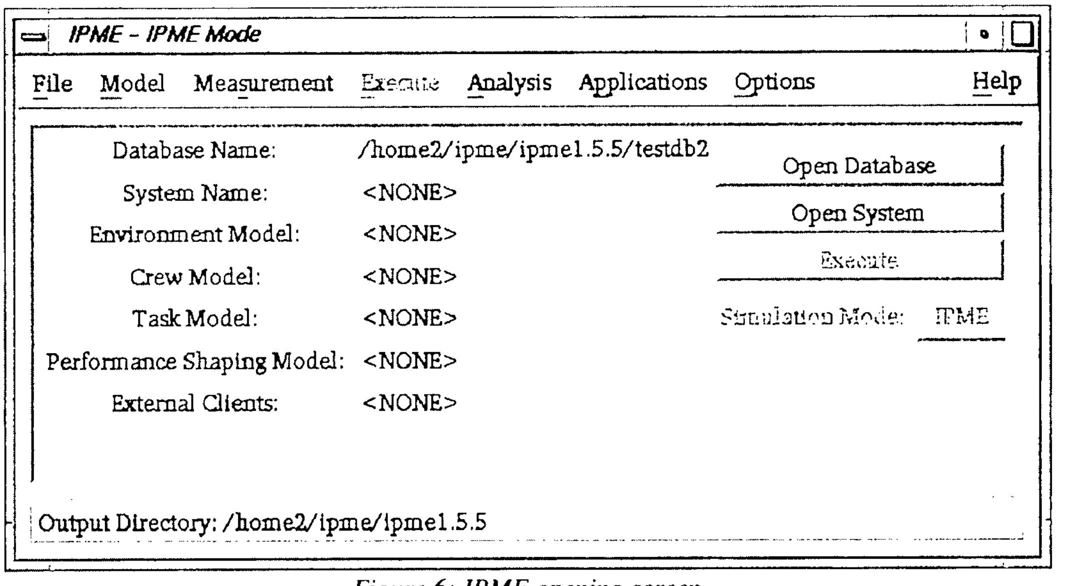


Figure 6: IPME opening screen

2. *Select system.* When the Open System button is selected, the System Description screen appears.

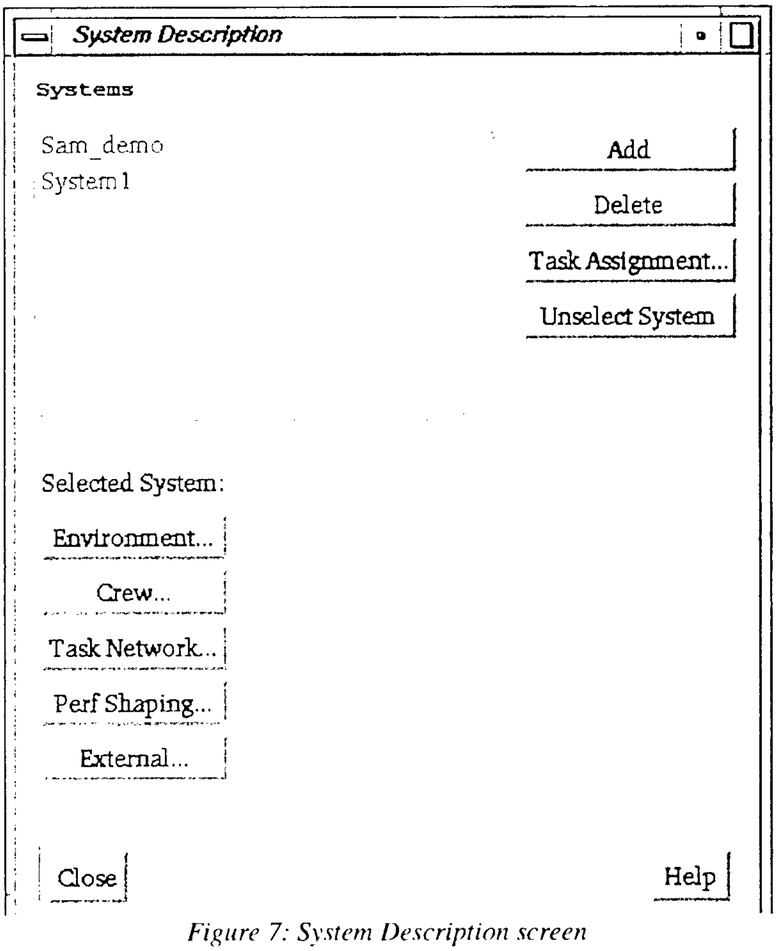
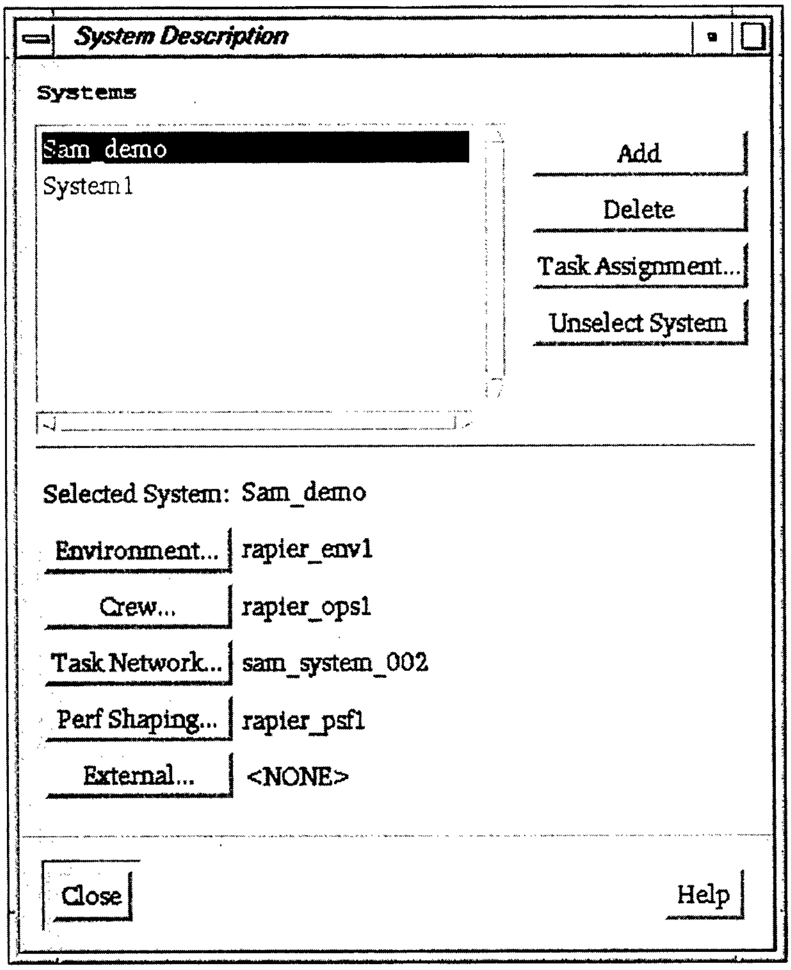


Figure 7: System Description screen

3. *Select system.* If the system is already available, click on the appropriate system in the list. The component models within that system are named, as shown in Figure 8.



4. When the select system screen is closed, the component model names are filled in on the opening screen.

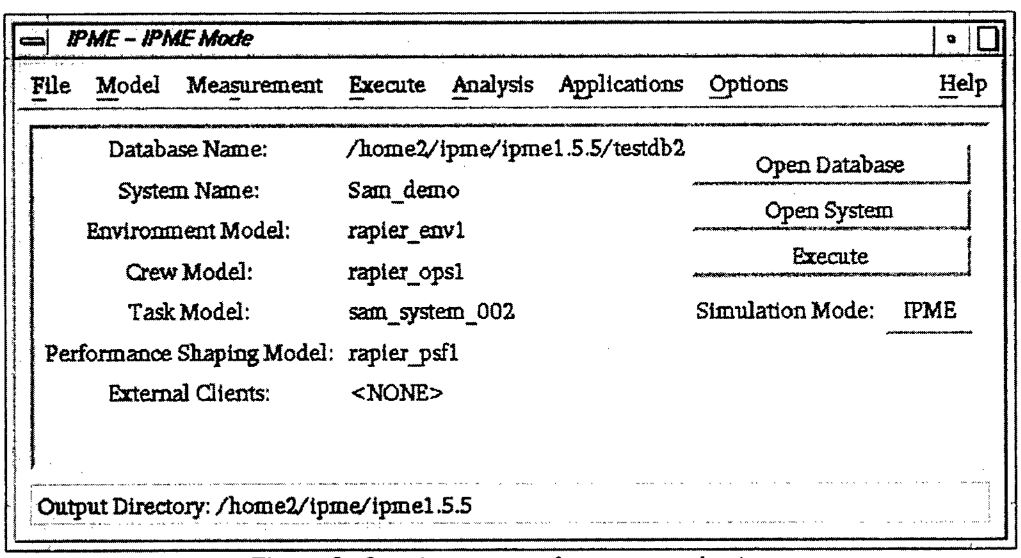


Figure 9: Opening screen after system selection

5. From the model menue on the opening screen, the task network model is selected, and a diagram of the task network is displayed as shown in Figure 10. The particular example is that shown in Figure 1.

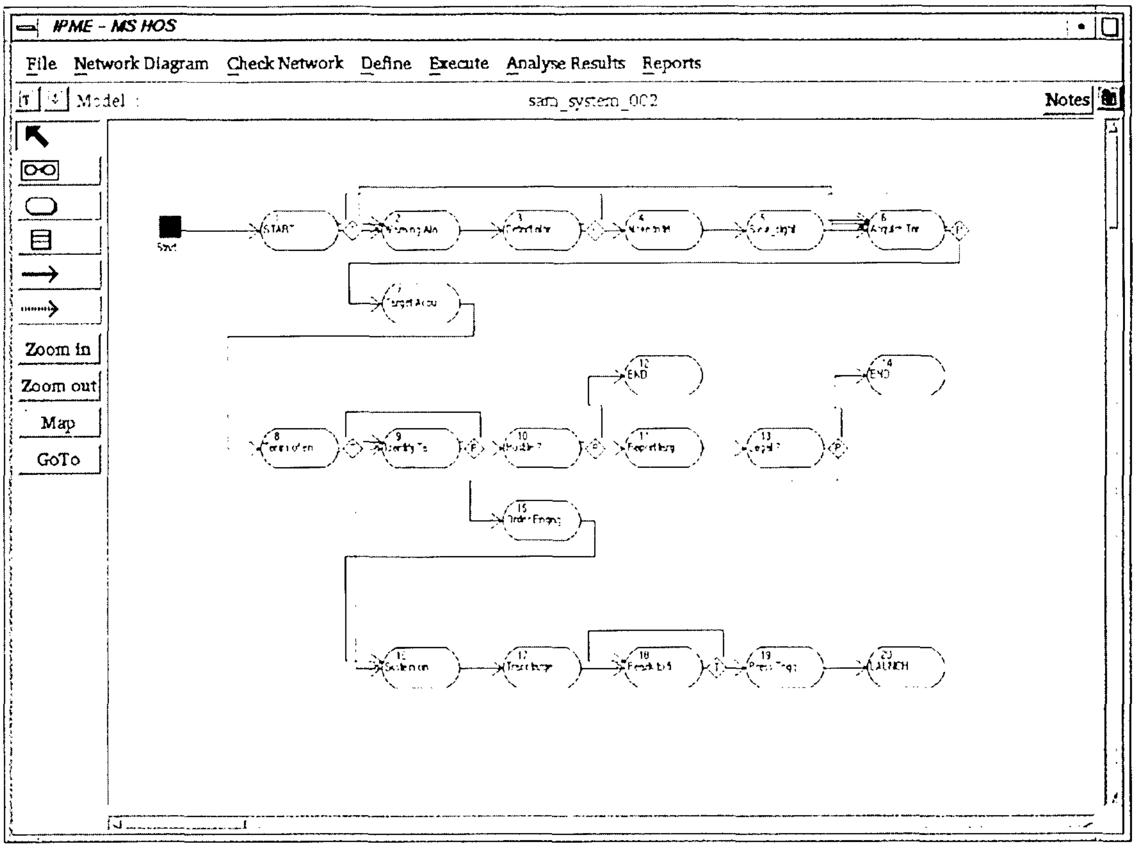


Figure 10: task network display.

6. The individual tasks are opened for editing by double-clicking on the task identifier. The screen for Task 3 - Warning Alarm is diaplayed in Figure 11. All the fields described in Section 1 are available for editing.

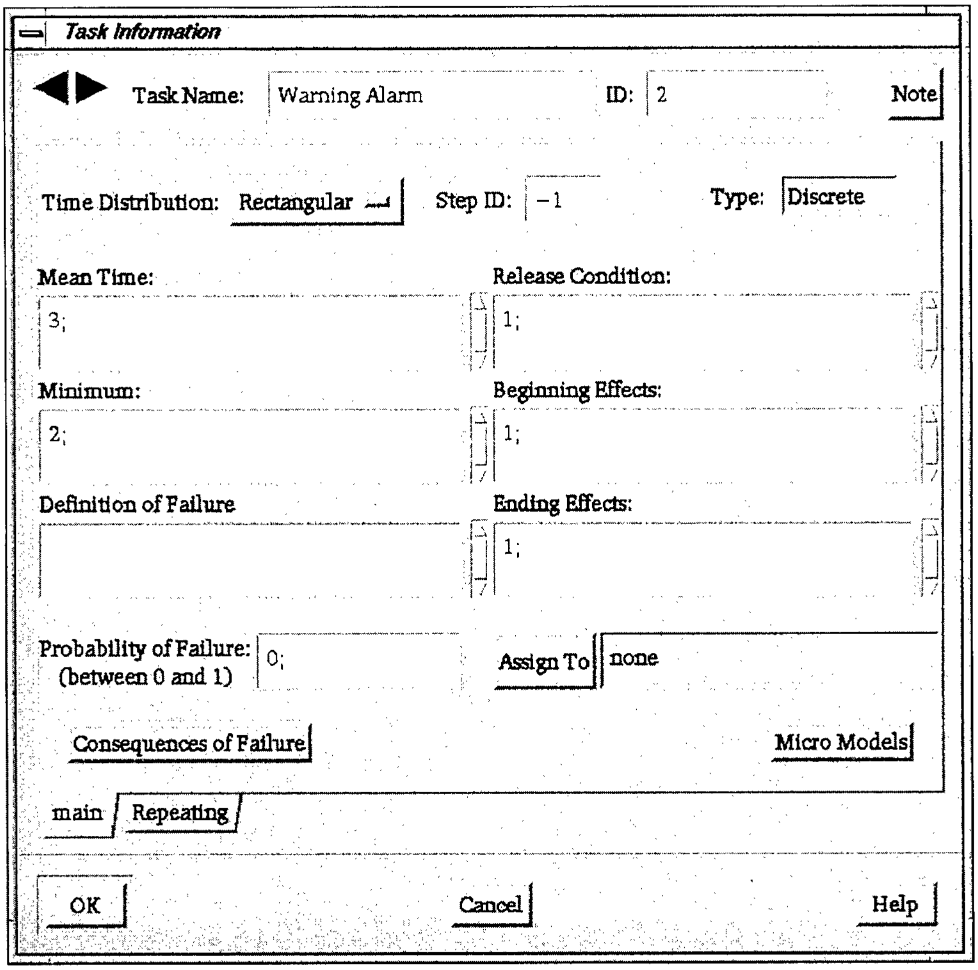


Figure 11: Task modification screen

7. There are a number of additional screens which are opened up by selecting 'Consequences of Failure', 'Repeating Task' or 'Assign to'. The most important of these, is the 'Assign To' button, through which the Operator who performs the task is allocated. As part of this dialogue, the taxonomic assignment allocation has to be made which determines the impact of the Performance Shaping Factors on the task. In addition, the assignment of values to the POP workload scales is made from this dialogue. The 'Assign To' screen is displayed in Figure 12.

The Operator can be assigned in one of three ways: Fixed (Static), expression - i.e. determined by some calculation, or 'Same as previous'. In the sample shown in Figure 12, the Operator is allocated statically to Commander.

Workload values have been assigned to the POP channels, and a taxonomic assignment has also been made.

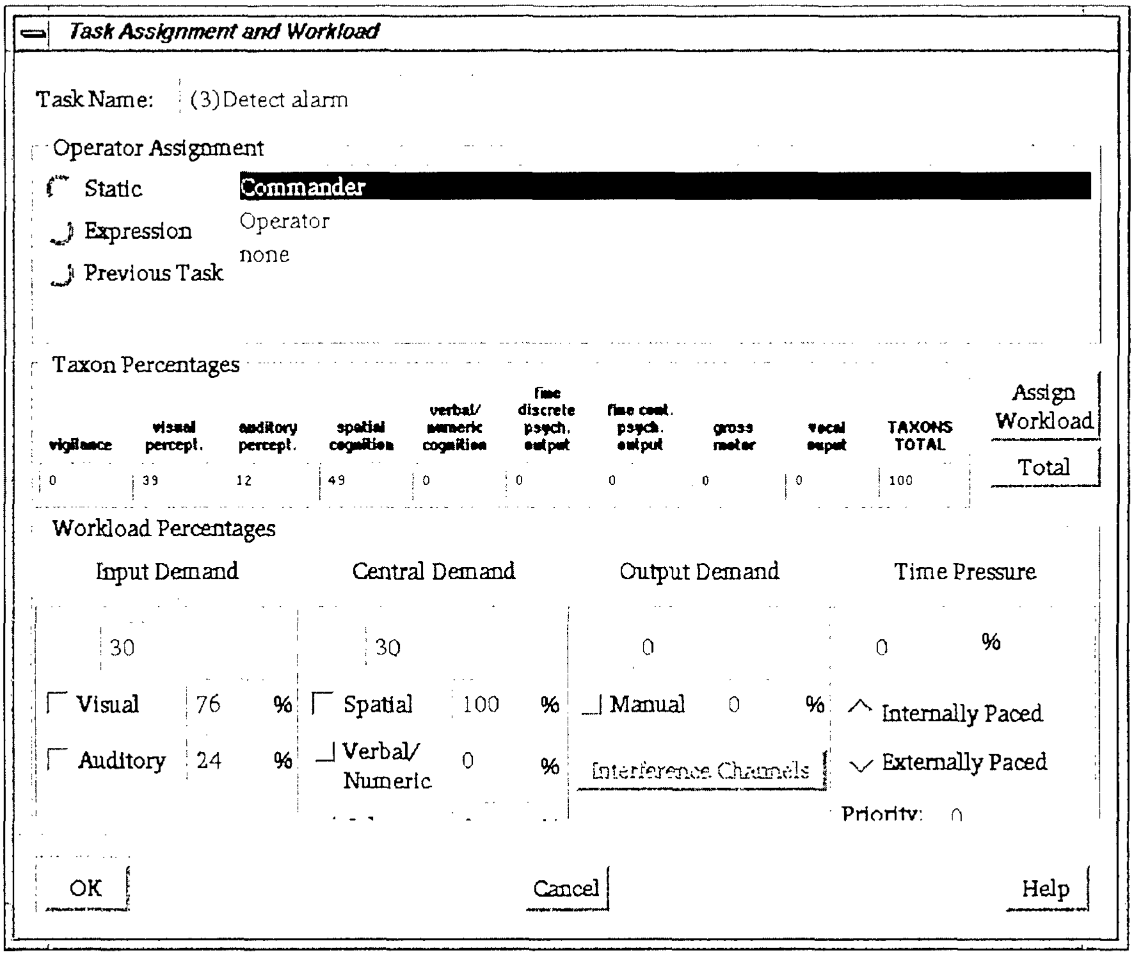


Figure 12: Operator assign screen.

8. The other component models can be edited in the same way as the task network. In Figure 13, the top level screen for the Crew model is displayed, showing two operators in the current crew.

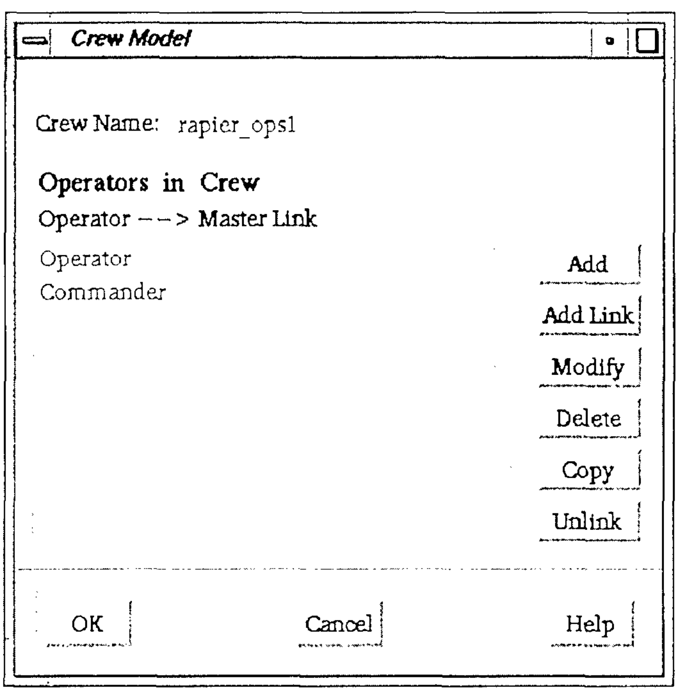


Figure 13: crew model top level screen

9. *Operator description.* There is a detailed description which can be filled in for each operator. In Figure 14, the top level screen is displayed for the commander. The associated Anthropometry screen is displayed in Figure 15.

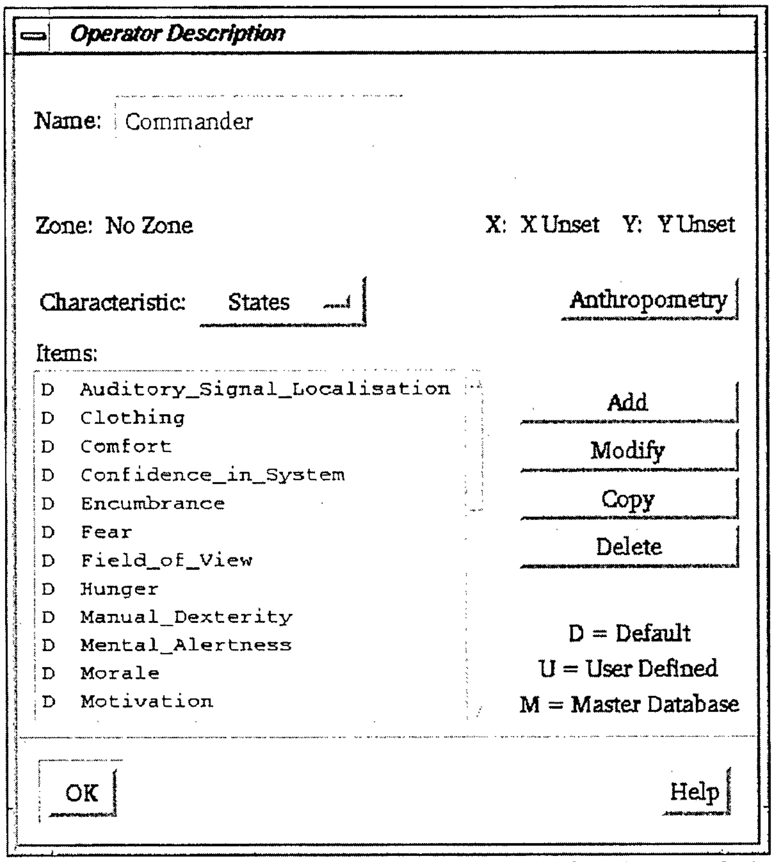


Figure 14: Top level Operator description screen for 'Commander'

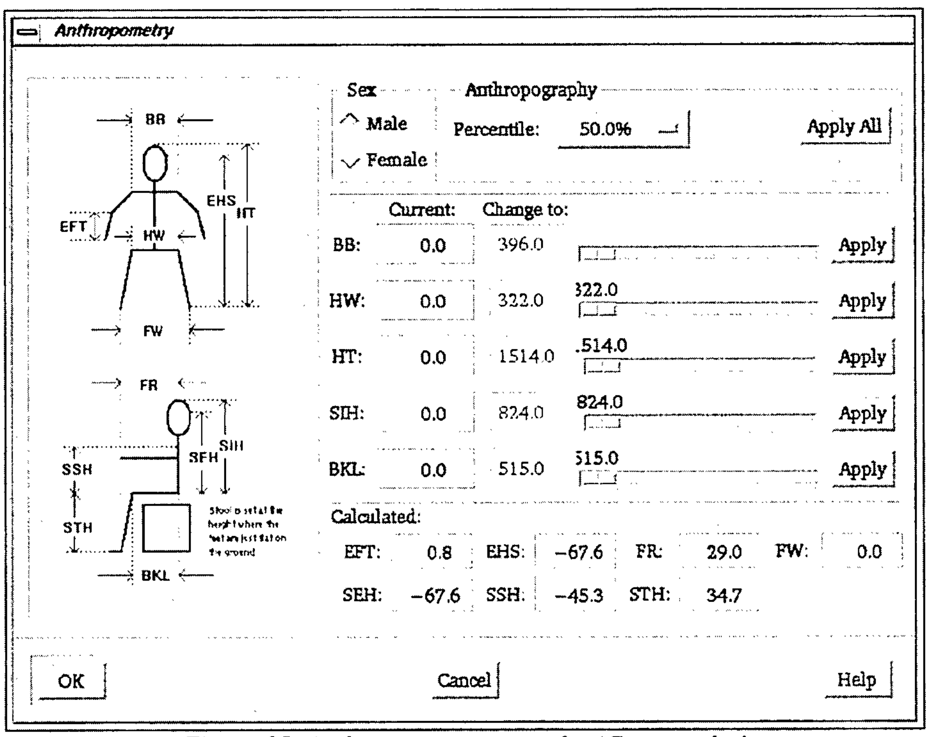


Figure 15: Anthropometry screen for 'Commander'

The full characteristics of an Operator are broken down into States, Traits and Properties. Each of these 'Variables' has a number of associated Attributes, and expressions which

determine the values of the attributes can be assigned as part of the simulation model. A typical application of this machinery is the definition of Mental_Alertness through Time_Since_Slept and Time_Of_Day as described in an earlier section.

10. *Environment model.* The Environment model contains 4 sub-sections: Physical, Crew, Mission and Threat. The top level screen for the Crew component of the Environment model is displayed in Figure 17. Each variable has both an initial value and an expression associated with it. The value and expression is modified by double-clicking on the appropriate variable.

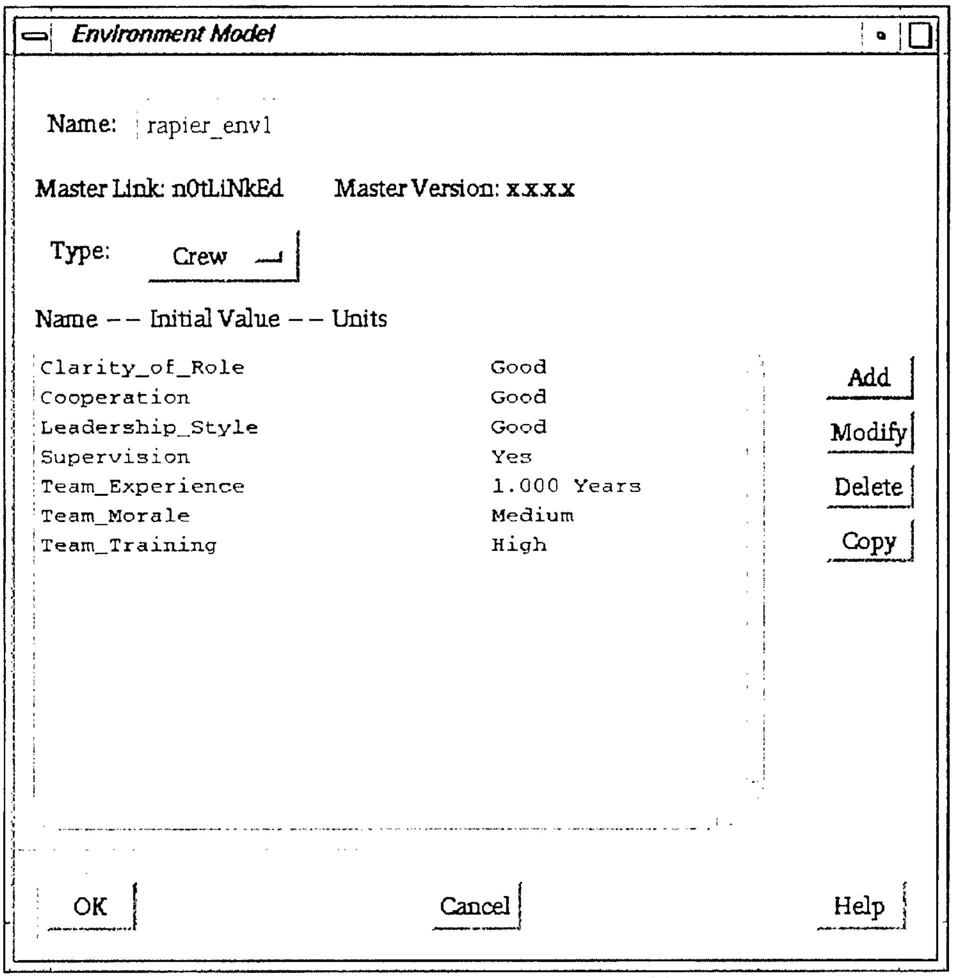


Figure 17: Crew environment variables

11. *Mission variables.* The mission variables screen is displayed in Figure 18.

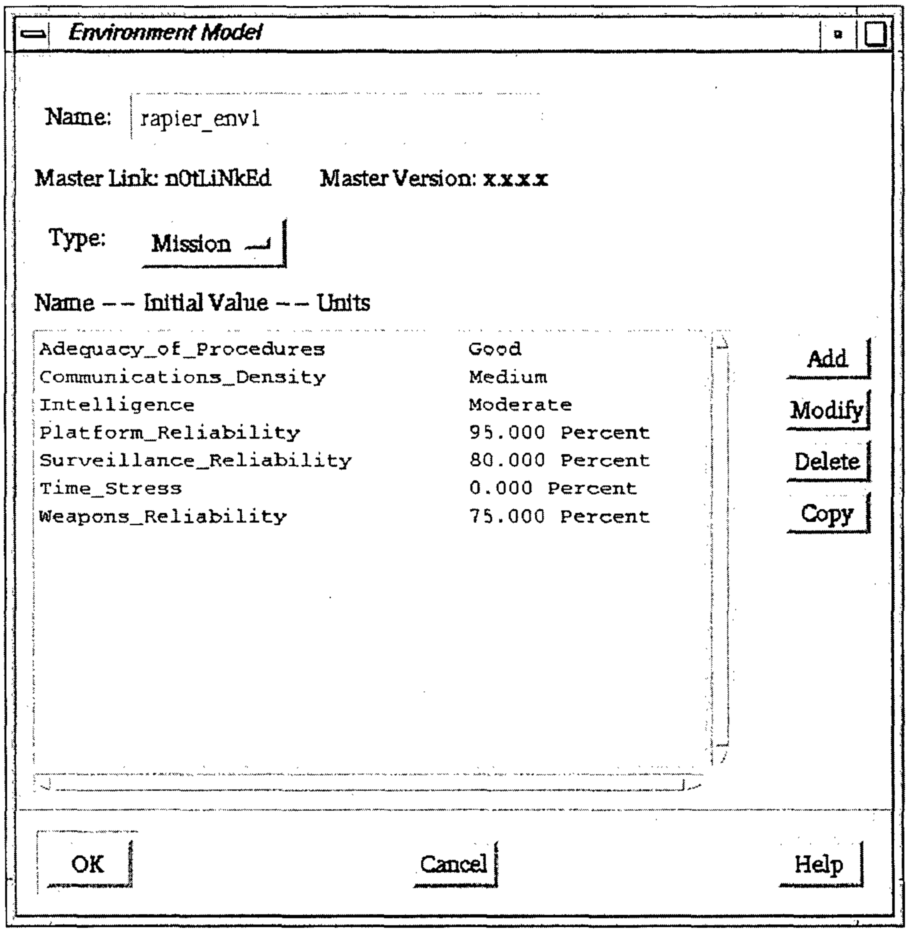


Figure 18: default mission variables

12. *PSF model.* The final model is the PSF model. The user selects and types in the expressions which form the Performance Shaping Model. Each individual Performance Shaping Function can be associated with a specific set of Taxons and Mean Task Time, Task error rate or be an intermediate function. The top level dialogue associated with the PSF model is displayed in Figure 19.

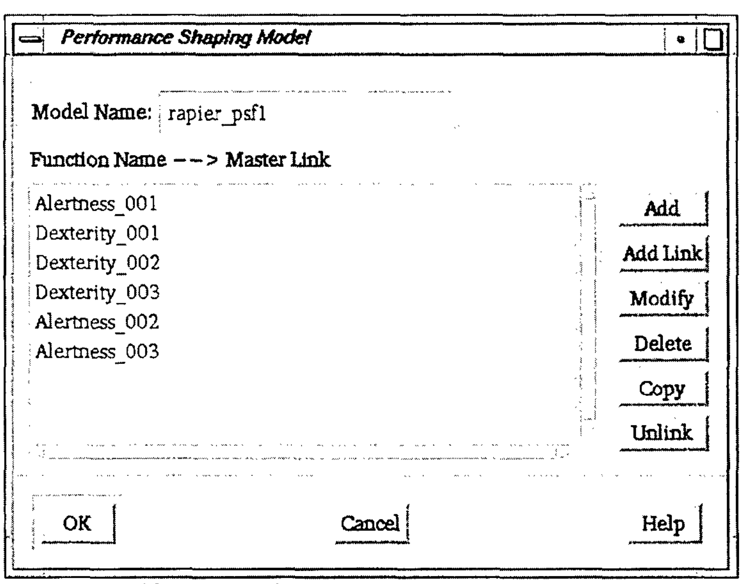


Figure 19: Top level Performance Shaping Model screen

13. *Performance Shaping Function.* To examine and modify the nature or action of an individual Performance Shaping Function, it is necessary to double click on the selected function, and the ancillary dialogue displayed in Figure 20, is opened.

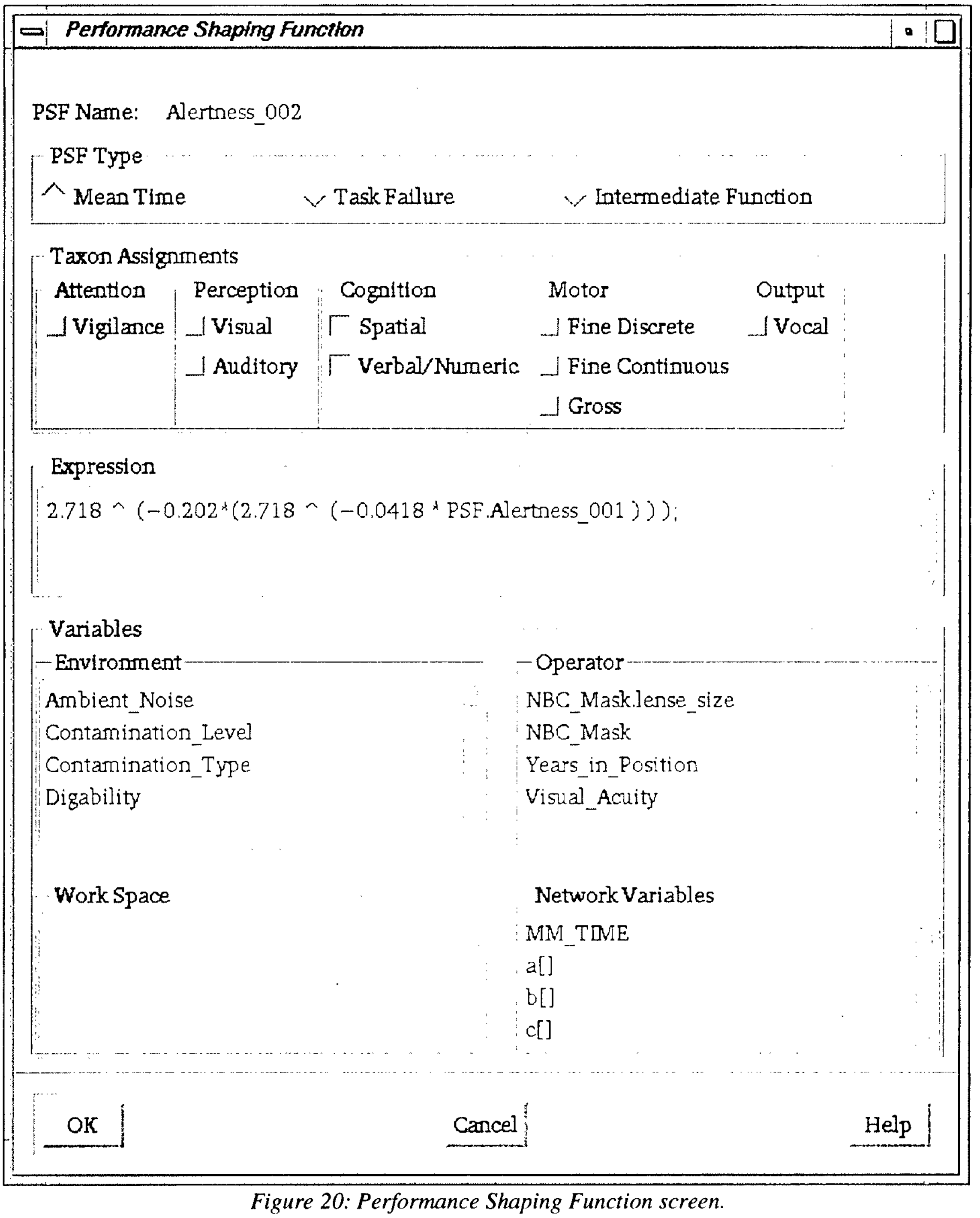


Figure 20: Performance Shaping Function screen.

This key screen consists of three parts: the nature of the function (Mean Time etc.), the Taxons on which the function acts, and the expression which is applied. The example shown in Figure 20 modifies the Mean time for Cognitive tasks, using the expression:

$$\exp(-0.202\exp(0.0418 * PSFAlertness\_001))$$

PSFAlertness_001 is an intermediate value calculated as part of the Performance Shaping Model; it is visible as the first Performance Shaping Function listed in the top level display (Figure 19).

When the model has been completed, it may then be executed. The execution options are selected from the Execution Settings screen displayed in Figure 21. There are a number of options which can be selected. during model testing, key options are "Display Variables" - which enables the user to track the value of variables as the simulation progresses - , and the animation option - which enables the user to follow the network logic flow.

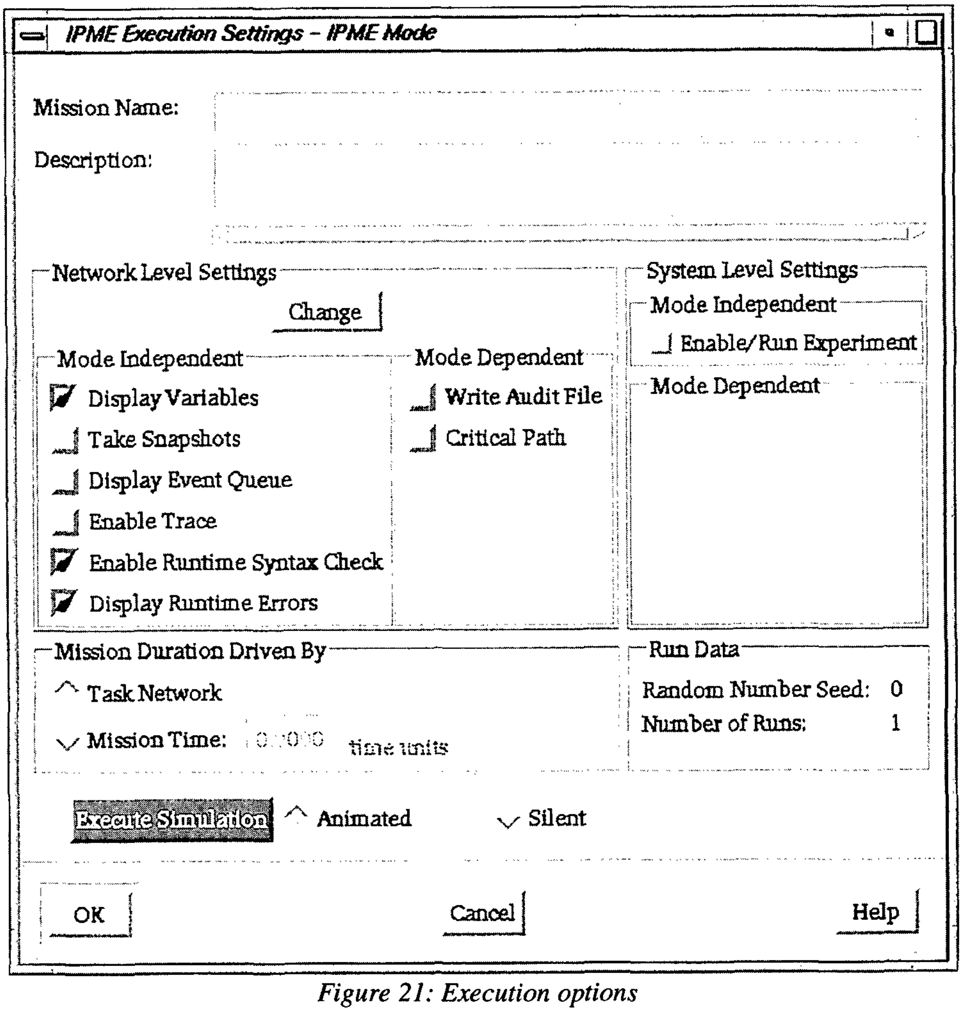


Figure 21: Execution options

When execution is started in Animated mode, with Display Variables turned on, two ancillary dialogues appear, as displayed in Figures 22 and 23.

The dialogue displayed in Figure 22 can be used to manage the execution of the simulation. The task network can be stepped event by event using the Pause / Step mechanism, or can be executed using the 'Start / resume' button. The speed with which the simulation executes can be controlled using the slider in the lower part of the dialogue.

The dialogue displayed in Figure 23 provides a display of the current value of selected variables at any stage of the simulation.

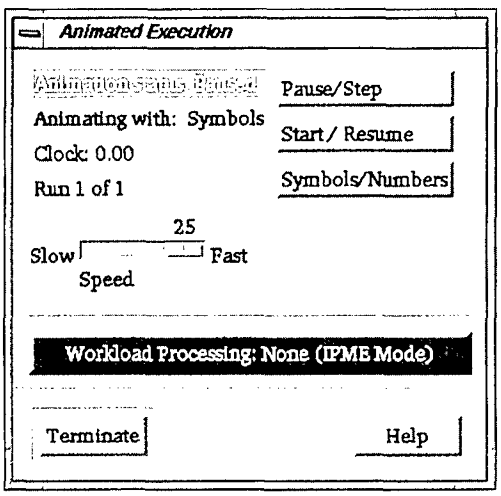


Figure 22: Execution control screen

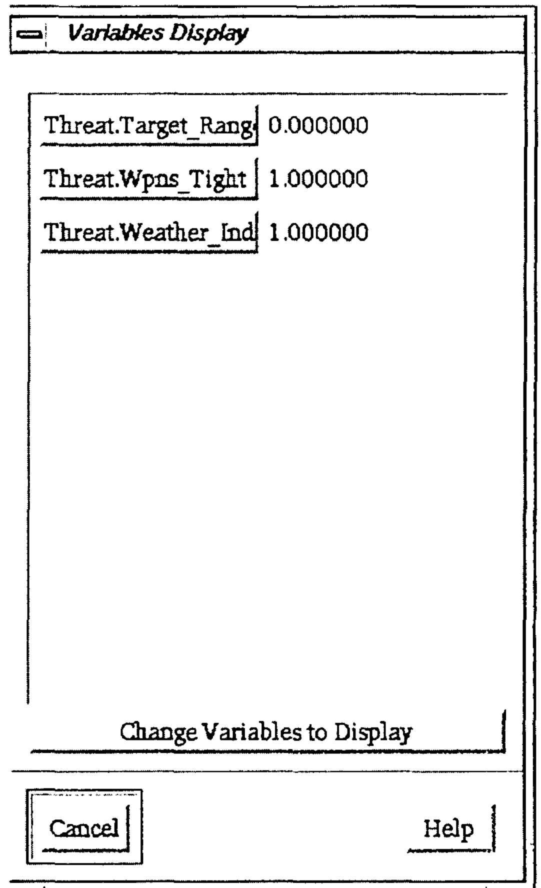


Figure 23: variable display screen

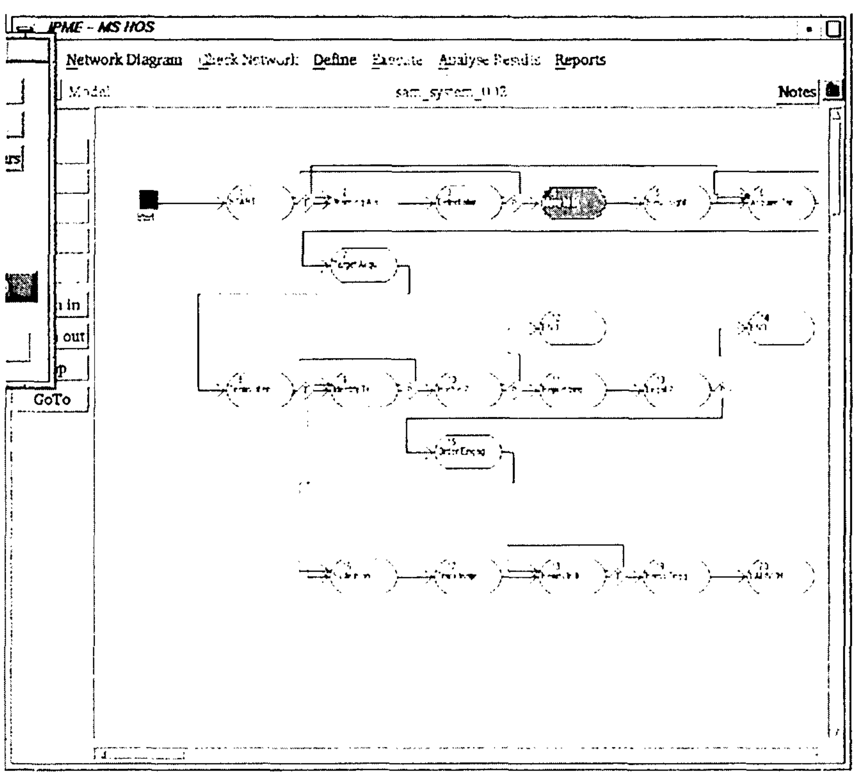


Figure 24: Network executing in animated mode

When a network is executing in animated mode, the currently executing task is shown in blue, as displayed in Figure 24.

Worked Example Of The Use of PUMA in a Function Allocation Task

Mr P Day
Roke Manor Research Ltd
Roke Manor,
Hants
UK

INTRODUCTION

The PUMA method and toolset was used in an allocation of function study, involving the re-engineering of a major civil Air Traffic Control system. As is the case in advanced, process-control like systems, one of the major issues facing designers is the extent to which functions formerly undertaken by humans in the system may usefully be automated. In the case of ATC systems, safety remains the paramount consideration, but there is also a growing requirement to increase system throughput as the levels of civil air traffic continue to grow. For this reason, civil aviation authorities around the world are increasing their level of investment in ATC systems, and in many cases replacing obsolete systems with new technology. ATC remains however a human-centred control activity, a situation that is unlikely to change in the foreseeable future, and hence one of the major issues that faces designers is the extent to which system functions may usefully be delegated to computer control while still keeping the human firmly in the loop.

The study described below was undertaken in this context, and is an illustration of the use of the PUMA method and toolset for the purposes of task analysis and workload estimation, thus enabling decisions on functional allocation to be taken.

The PUMA Method

The basic PUMA method involves a number of stages:

- Establishing a base-line of controller activities by analysing (or drawing upon a pre-existing analysis of) ATC activities as they are currently performed;
- Breaking those activities down into those fundamental components which impose a predictable loading on the controller;
- Establishing what new circumstances or procedures are to be examined using the toolset, which might for instance involve introducing changes to the fine task structure (typically associated with the use of new computerised support tools), and then setting that in the context of a scenario of aircraft movements within a sector;
- Calculating workload, using a technique based on Wickens' "multiple resource theory". This involves the concept of multiple channels within the user, upon which demands are made when tasks are undertaken, and which may conflict when complex tasks are carried out.

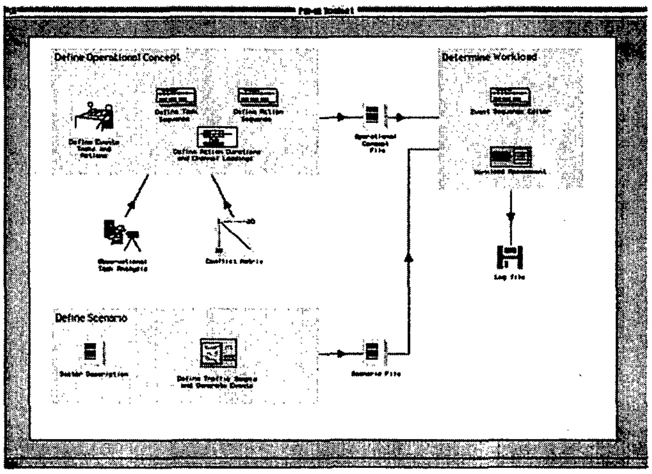


Figure 1 PUMA Top Level Diagram

The PUMA method is supported by the PUMA toolset, which has been built on top of the pre-existing NMSE (Network Modelling Support Environment) software, a LISP-based, object-oriented model-builder. The PUMA toolset consists of a family of independent tools with a common "look and feel", and the ability to exchange data between them readily. The philosophy has been followed that any data file is stored in a human- readable, English language ASCII form, and can be edited either within the tool that created it, or in text form within any standard word processor.

ANALYSIS

The starting point for the use of the PUMA method is a Definition of the Operational Concept, that is a

process of defining and linking together the roles, tasks, events and actions involved in the area of ATC under study. A "role" may be seen as being associated with the performance of particular duties. In the ATC context different controller roles exist and it is necessary to be able to associate a person with a particular task. In the PUMA method, a "task" consists of a number of "actions". The granularity of these is such that an action places an unvarying demand on the user's cognitive processing channels, while a task - which probably consists of a number of actions, some overlapping - is a recognisable ATC duty, such as giving an aircraft a clearance, or accepting a new aircraft into the sector. An "event" is an externally generated phenomenon that causes a controller to take some action (be it an overt, observable action, or covert, cognitive action).

In defining an operational concept one draws upon many sources, including the formal procedures that controllers are taught, the published literature, and the descriptions given by controllers themselves of their work. When PUMA is being used to examine the workload implications of a new way of working, the operational concept will typically differ in various ways from the baseline that has been established by these means.

The Membership Editor (ME) allows the user to represent in list form the roles, events, tasks and actions, and define and display the links between them. All the information concerning events, tasks and actions is stored in the Operational Concept File, which can be edited textually using the Operational Concept File Editor. The toolset is constructed so that the various editors both write information to, and draw information from, the OC file as necessary.

The definition of the operational concept will typically have built upon an initial Observational Task Analysis (OTA), and so it was on this case. This involved observing and videotaping controllers performing their duties (with their permission), and then relating observed actions to tasks. For the system under study, a range of controller positions were studied, covering tower, TMA (Terminal Manoeuvring Area - the airspace around an airport, where aircraft are climbing and descending), and en route control. In every case the observational sessions were preceded by interviews with controllers who understood the airspace and the controller tasks we would be observing, so that the team undertaking the study were fully primed and able to understand what would be happening. Observational sessions were scheduled to take place during the busy times of the day. The OTA approach used in PUMA involves having the controller talk through the videotape of his actions immediately after his spell on duty has ended, thereby allowing a good insight into not just what he did (and did not do), but why. These interviews were in turn recorded on video. These recordings, and the direct video recordings of the controller doing his task, enabled a subsequent video analysis which resulted in the expression of the operator's activities in terms of actions, and the time and duration of these actions.

The OTA Support Tool (OTAST) allows the graphical representation of overt and covert actions against a timeline. It also allows the user to define the tasks that the controller was undertaking, and to associate the component actions with an appropriate task. Thus, a task contains a number of actions. This "grouping" of actions with tasks is supported by the Task Structuring Tool (TST), which is embedded within the OTA Support Tool. Task structuring involves analysing the actions obtained from the OTA, and interpreting them in the light of the knowledge of what the controller was doing. Actions are of a granularity such that the demands on the controller's information processing channels are constant throughout the conduct of that action. Tasks may involve the execution of a number of actions, (which may themselves overlap), but are reasonably consistent in the actions they contain. Tasks are of a granularity such that they may be edited and re-ordered by controllers or other ATC-knowledgeable people when creating new scenarios. Thus it was with the ATC analysis undertaken that after the analysis and grouping activities, sessions were held with domain experts (the controllers who had given advice on the airspace, tasks etc. before the observations took place). They verified that the analysis was sound, and that the tasks identified and the actions they contained belonged together. In some instances, they were able to point out small errors in the analysis.

One useful feature of the PUMA toolset is the support it provides for the video analysis process. Traditional video analysis is done with a video recorder, shuttling it back and forth over the sequence of interest, and using 'freeze-frame'. PUMA provides special support for the video analysis process, in that it allows the user to select a video sequence on tape, capture it onto hard disk (in a standard compressed format), and then link it in to the actions and tasks being analysed. The user can then open up a video window within the OTAST, and play the video sequence of interest complete with sound. Since the video data is now coming from hard disk, the user can rapidly scroll up and down the sequence, pause it, inch forwards or backwards a frame at a time, and so on. The video

is fully integrated, so that as it plays a vertical line scrolls across the OTAST, indicating the relevant actions and tasks. Similarly, dragging the vertical line along the task sequence moves the video clip to that point. Correcting the OTA data is easily done, by selecting a task or action, moving the video to that point, then clicking a button to correct the start or end time. Multiple video windows may be opened and run within the OTAST, for instance to see different instances of the same task being performed.

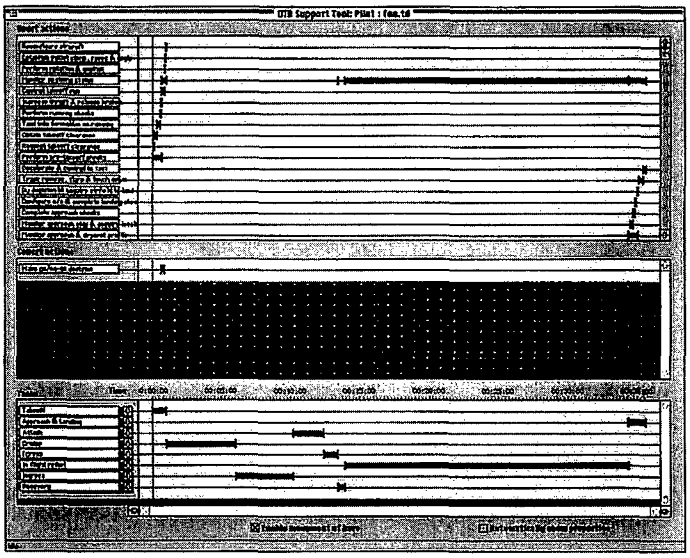


Figure 2 Operational Task Analysis Support Tool

Task "generification" (a way of deriving an average or generic version of the observed tasks) was then undertaken using the Task Generification Tool (TGT), which is embedded within the OTA Support Tool. When this is invoked an automated process examines all the instances of each task identified during the OTA, and determines how internally consistent the task is in terms of the actions it contains, their length, and the overall length of the task. The generic version of the task is then generated, along with the plus-one-standard-deviation version. Naturally, it is important that the user examines the generic version of a task and edits it as necessary, since like the "average" family of 2.4 children, it might not make sense in a single instance. (A concrete example occurs with ATC speech. In most cases, conversations occur on a turn-taking basis, and each observed instance of the task might reflect this. A generified version, however, might contain overlaps, that would when put through the workload calculation algorithm show unrealistic workload peaks).

Thus the OTAST and its embedded tools, the TST and TGT, allowed the creation of a baseline definition of tasks and actions within the roles studied. PUMA also supports the creation of tasks from a conceptual level rather than simply from observation, and this is done using the Membership Editor, which allows the user to re-define the nature of the tasks to be performed in a top-down style. From the membership editor the user can call two further editors, which allow the operational concept to be explored from two different perspectives.

The Task-Action Ordering Editor (TAO Editor), called from the Membership Editor, allows the user to look at each task that has been defined, see the actions within it (both overt and covert), edit the durations and channel loadings of those actions, and calculate the workload for each of those tasks. Furthermore, the TAO Editor allows the user to see at a glance which role is connected with a task. A further feature of the TAO Editor is the ability to select all actions, and edit their durations and channel loadings. When the changes made using the TAO Editor are saved (to the Operational Concept file), they form the new global definitions of those variables.

The Event-Task Ordering Editor (ETO Editor), also called from the Membership Editor, allows the user to look at tasks from the perspective of events, i.e., the external triggers. It also allows the user to see which roles are associated with those tasks, and to calculate the resulting workload for that role. The start time of the events can be edited using the ETO editor. When the changes made using the ETO Editor are saved (to the Operational Concept file), they form the new global definitions of those variables.

Thus the new editors allowed a range of operational concepts to be defined very readily, and then examined from different perspectives, in terms of the workload involved in tasks, and the workload associated with individual roles. In addition, the use of a single master file that defines everything to do with the activities of the controllers (the Operational Concept file) made it easy to maintain configuration control of alternative operational concepts. The Membership Editor incorporates a built-in report generation facility, which was used to automatically create detailed reports of the operational concepts.

Having undertaken the activities outlined above, the next step was to develop the scenarios for which workload was to be calculated. The Scenario Builder/Editor (SBE) tool supports the process of creating an ATC scenario, which would typically involve defining a sector of particular dimensions, with reporting points, standard routes, and a number of aircraft of identified types with realistic flight

plans. The SBE gives the user a graphical representation of the area of interest, rather like a modern colour radar display, with the ability to zoom in on any area of interest. The user can then exercise the scenario, with the aircraft flying according to the flight plans in the system. He can build up the scenario from an ATC point of view by adding extra tasks to the scenario as he wishes (such as having the controller put an aircraft on a radar heading, or requesting an aircraft to climb to a new flight level), and this is then logged into the scenario file as he progresses. The full list of ATC tasks that he can create for the controller is a function of the earlier analyses undertaken. The SBE has certain aircraft-related tasks built in (climb, descend, adopt heading, resume own navigation to a beacon, etc.), and can read in the Operational Concept file giving the tasks which do not directly affect the display of aircraft movements, but do nevertheless affect controller workload, and hence must be part of the scenario.

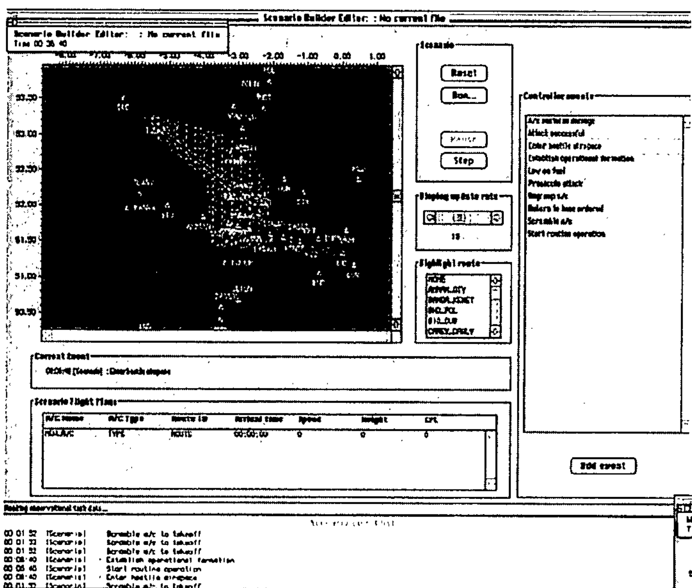


Figure 3 Scenario Builder Editor

The user may continue building (and editing) the scenario until it represents what he wishes, and then he saves the Scenario File for subsequent execution and workload calculation by the toolset. While the SBE provides full support to the user in generating sectors, flight plans, reporting points, and events, it is unlikely that users would want to build these from scratch every time, but would rather prefer to call up existing scenarios from file and modify them as necessary. This modification can be done graphically using the tool, or by directly editing the Scenario File using the Scenario File Editor. Within the file all the parameters are expressed in plain English text, and may be edited accordingly. Thus it was that in the current study use was made of pre-existing definitions of operational scenarios, based on predictions of traffic levels in future time frames, and at particular times of the day.

The next step involved the user reviewing and editing the complete task sequences based on the OTA (the baseline), modified in the light of future operational concepts. Support for this process was provided by the Event Sequence Editor (ESE) tool, which provides a graphical display of aircraft movements (as with the SBE), but as it plays out the aircraft movements it also displays the various controller tasks as timelines. By this means, the team was able to gain the best possible understanding of the scenarios and the controller tasks within it.

Finally, when the complete process of scenario and task editing had been completed, it was possible to invoke the Workload Assessment Tool (WAT) (which is embedded within the ESE), and again play each scenario through, this time also observing the curve of workload against time. The WAT also allowed the team to see the workload data expressed in a histogram form, with the amounts of time spent at each workload level being displayed graphically. (All the PUMA tools which generate graphs can write data out in a format usable by most spreadsheets and charting packages). The WAT has a batch-file mechanism which supports unattended multiple runs with different scenarios and/or operational concepts, each being logged in different ways if necessary. It also provides a comprehensive set of data logging facilities, to allow the data produced during a run to be recorded and analysed further using other packages.

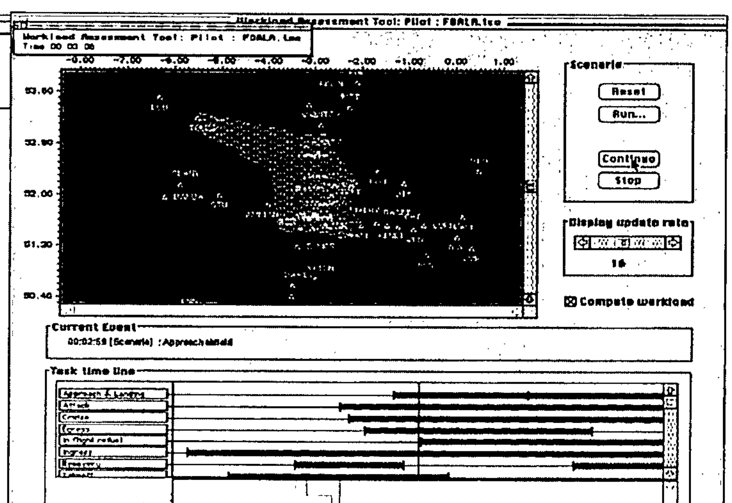


Fig 4 Workload Assessment Tool

CONCLUSION

The process described above was undertaken for controller tasks throughout the ATC system. In each case the process of establishing baseline tasks from observation, creating modified versions of those tasks, and then exercising them within the context of projected future scenarios of aircraft movements, provided an extraordinarily valuable insight into the potential value of those task modifications. In this context, PUMA cannot be a full substitute for high-fidelity man-in-the-loop simulations, but it may be seen as a most cost-effective way of cutting down the search space for potential solutions to complex future system design questions.

Application of an Anthropometric Tool to Cockpit Layout

Gretchen Burrett
Defence Evaluation and Research Agency (DERA)
Aircraft Test and Evaluation, Boscombe Down
Salisbury, Wiltshire, SP4 0JF
UK
Gretchen@gregory-harland.co.uk

1. SUMMARY

Anthropometric tools are used to assess human interaction with workplace layout in terms fit, reach, and vision. As humans do not come in a standard size, these tools address the range of potential users, from very small to very large. This paper provides an example of how Anthropometric tools can be used to help optimise cockpit layout. Jack® is used as an example tool.

2. PROBLEM SPACE

The physical characteristic of pilots must be considered to achieve effective human performance in the cockpit. Failure to consider aspects such as the range of body sizes (Anthropometry) of pilots or their physical strength can result in a wide variety of problems, including the following:

- Serious injury during ejection, for example, injuries resulting from collision of legs with display surfaces.
- Inappropriate force available to apply to breaks during landing because rudder pedals are placed too far away.
- Errors operating Hands on Throttle and Stick (HOTAS) controls because small hands cannot adequately reach finger operated controls.
- Inability to read head up and head down displays if the seat cannot be adjusted to allow pilots with a particularly short or tall sitting height to position themselves at the appropriate angle.
  Collision between helmet-mounted displays and the canopy for tall pilot, which restricts head movement and can negatively impact visual tracking of enemy (or friendly) aircraft.

These problems can be overcome by the application of one of several tools currently available which allow designers to evaluate the impacts of various design concepts on the anticipated user population. These tools can also be used to assess maintainer tasks.

3. PROCESS

The generalised process for performing an Anthropometric assessment is described in Figure 1. The stages are expanded upon below, using Jack® as an example tool.

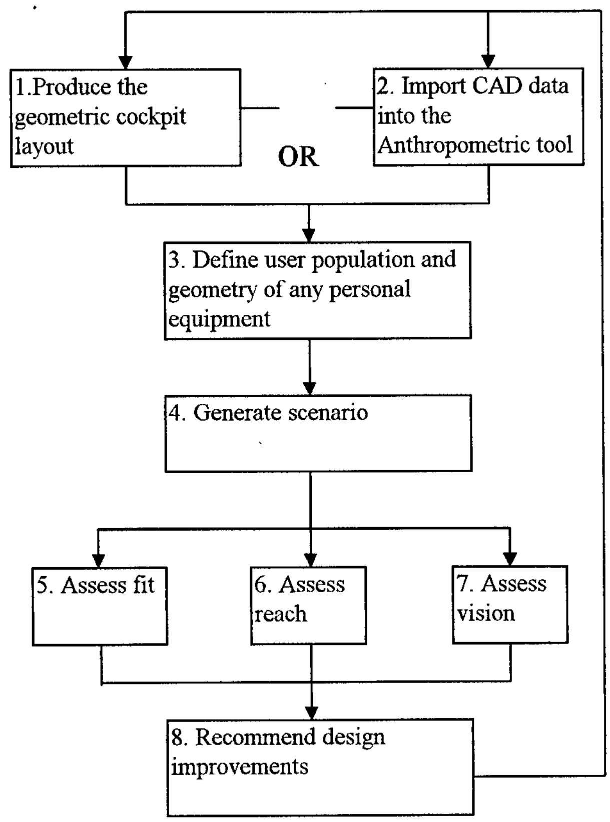


Figure 1: Generalised Anthropometric assessment process

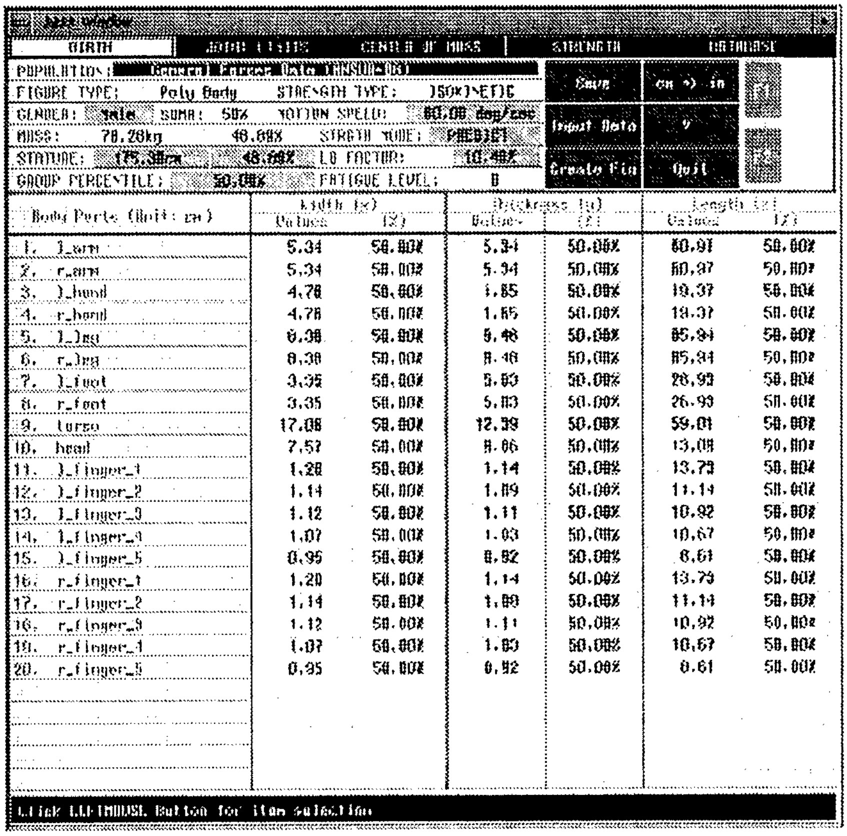
i. **Produce the geometric cockpit layout.**
If a CAD representation of the cockpit does not exist, the first stage is to generate a three dimensional representation of the cockpit. Jack® provides basic CAD functionality, and allows production of a workplace from a series of geometric shapes. This can be achieved using conventional CAD techniques. Using a combination of pop-up

menus and point-and-click, Jack® the user to specify the movement of items within the cockpit such as yokes or seats. The range of movement of these items can have a significant impact on fit, reach and vision. Therefore, they should generally be defined in the model. It is also important to specify the geometry of any personnel equipment such as helmets or parachutes which can be attached to the Jack® figure for a more realistic assessment.

ii. Import CAD data into Jack®.
To save effort, an existing CAD representation can be imported into Jack® using the import facility. Some CAD formats can be imported directly, others require the use of a conversion program.

It will then be necessary to input, using point-and-click, the range of movement of various cockpit items and the dimensions of personnel equipment, as discussed above.

iii. Define user population. Jack® comes equipped with default data on the US military population. Using he Spreadsheet Anthropometric Scaling System (see figure 2), it is easy to alter this data. This form is also used to select the size of the human figure(s) displayed within the CAD environment. This is achieved by selecting the desired percentile (usually approximately the 5th percentile represents a very small person and the 95th percentile represents a very large person).



Body Parts (Unit: cm)	Width (x) Values	(%)	Thickness (y) Values	(%)	Length (z) Values	(%)
1. l_arm	5.34	50.00%	5.34	50.00%	60.97	50.00%
2. r_arm	5.34	50.00%	5.34	50.00%	60.97	50.00%
3. l_hand	4.78	50.00%	1.65	50.00%	18.37	50.00%
4. r_hand	4.78	50.00%	1.65	50.00%	18.37	50.00%
5. l_leg	8.38	50.00%	8.48	50.00%	85.94	50.00%
6. r_leg	8.38	50.00%	8.48	50.00%	85.94	50.00%
7. l_foot	3.35	50.00%	5.03	50.00%	26.93	50.00%
8. r_foot	3.35	50.00%	5.03	50.00%	26.93	50.00%
9. torso	17.08	50.00%	12.39	50.00%	59.01	50.00%
10. head	7.57	50.00%	8.06	50.00%	13.08	50.00%
11. l_finger_1	1.20	50.00%	1.14	50.00%	13.73	50.00%
12. l_finger_2	1.14	50.00%	1.09	50.00%	11.14	50.00%
13. l_finger_3	1.12	50.00%	1.11	50.00%	10.92	50.00%
14. l_finger_4	1.07	50.00%	1.03	50.00%	10.67	50.00%
15. l_finger_5	0.95	50.00%	0.92	50.00%	8.61	50.00%
16. r_finger_1	1.20	50.00%	1.14	50.00%	13.73	50.00%
17. r_finger_2	1.14	50.00%	1.09	50.00%	11.14	50.00%
18. r_finger_3	1.12	50.00%	1.11	50.00%	10.92	50.00%
19. r_finger_4	1.07	50.00%	1.03	50.00%	10.67	50.00%
20. r_finger_5	0.95	50.00%	0.92	50.00%	8.61	50.00%

Figure 2: Screen print from Jack® showing the Spreadsheet Anthropometric Scaling System.

iv. Create Scenario. The Jack® figure can be animated to follow a sequence of movements. Special features such as reverse kinematics allow the figure to maintain realistic motion and to maintain balance. Scenarios are created interactively by point-and-click to specify motion, and via an interactive timeline.

v. Assess fit. Once the Jack® figure has been inserted into the CAD environment and a scenario has been created, it is a simple matter to assess fit. The pop-up menus are used to turn on the collision feature. During a scenario run, any time the Jack® figure collides with an obstacle such as the yoke or canopy, the obstacle is highlighted in red. (See figure 3.)

Figure 3: Screen print from Jack® illustrating a collision.

vi. Assess reach.
Reach can be assessed within Jack® by using the ruler feature. Once the feature is activated using the pop-up menus, a ruler will appear showing the additional distance to go in reaching any item in the scenario where the Jack® figure's reach falls short. (See figure 4.) It is possible, using the pop-up menus to specify type of reach required (e.g., touch or grip) as well as the amount of body motion allowed (e.g. can or cannot lean forward at the waist).

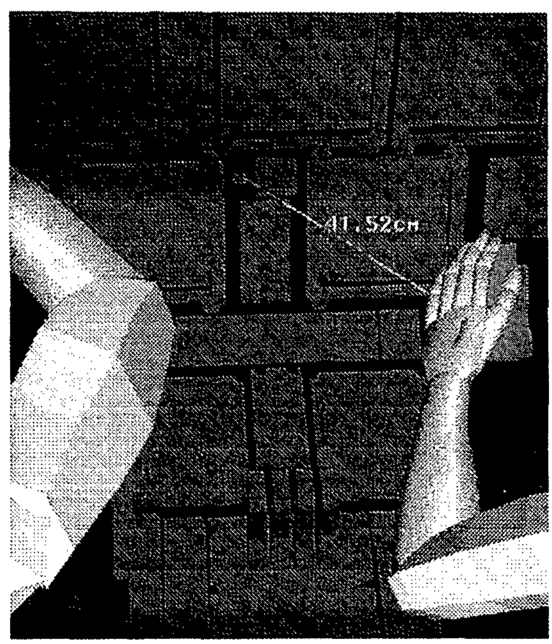


Figure 4: Screen print from Jack® illustrating reach ruler.

vii.Assess vision.

There are two ways to assess vision. One way is to bring up a window showing the eye view during the scenario. As the Jack® figure moves his/her head, the eye view perspective is displayed in the window. The second method is turn on the translucent view cone facility. This illustrates what the figure could see without eye movement, or with eye movement but no head movement. The view cone can be overlaid on the eye view perspective to give a better indication of where critical displays should be placed. This is illustrated in figure 5.

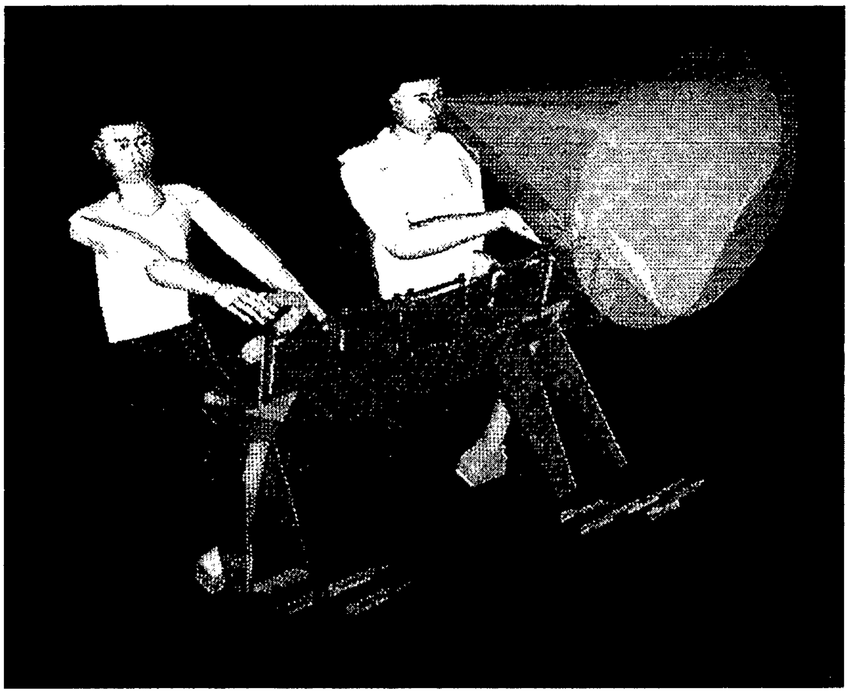


Figure 5: Screen print from Jack® showing vision envelope.

viii.Recommend design improvements.

The final step is to use the information obtained through the Anthropometric assessment in formulating recommendations for design improvements. The model provides a visually compelling argument which graphically illustrates any problems with fit, reach or vision for the range of expected users. It can also be used to assess proposed changes to the design.

4. Limitations

Compatibility of tools with the range of existing CAD packages is less than perfect. Because no common data format exists for CAD, some CAD files are more compatible with the Anthropometric tools than others. In general, a conversion file can be written for most file formats.

Design layout and physical dimensions must be available.

Some tools use joint-to-joint measurements, rather than anthropometric measures (e.g., top of knee to bottom of foot) which are more frequently available in the published literature. This can make interpolation to new populations more difficult.

5. Facilities Resource Requirements

Depending on the type of Anthropometric tool selected, these range from a PC to a Silicon Graphics machine. Jack®, which can import CAD files, and requires at least an Indigo 2 Silicon Graphics engine. Jack® the trademark of the University of Pennsylvania, where it was created by the Centre for Computer Graphics Research funded by US military and commercial customers. It is distributed by Transcom Ltd.:

www:\transcom.com

Human Reliability Assessment Tools - PHRASE 2

Brian Buck and Gretchen Burrett
GEC-Marconi Research Centre,
West Hanningfield Road,
Great Baddow,
Chelmsford,
Essex, CM2 8HN
UK

1. SUMMARY

Human Reliability Assessment (HRA) tools seek to quantify the likelihood of human error given that error mechanisms have been identified. They form an integral part of a larger process of Human Reliability Assessment, see Figure 1. HRA has traditionally been used primarily in the process control industries, but some methods are appropriate to military applications. Its use requires skilled practitioners.

HRA is not a substitute for detailed human factors assessment when the objective is to maximise human performance. However, it will assist in directing design and evaluation effort where the human contribution is most critical. This paper outlines how HRA tools can be applied to cockpit design and describes the HRA process. PHRASE 2 is used as an example tool.

2. PROBLEM SPACE

The pilot's interaction with cockpit equipment contributes to the effectiveness and safety of the cockpit system. Human Reliability Assessment (HRA) tools can be used to specify both the types of human error that are likely to occur and the probabilities associated with these errors. The HRA process is intended to predict only *gross* differences in human performance, when there are several ways to achieve mission success using permutations of human tasks and technology.

HRA has two principal applications, in each case, the human contribution is potentially critical to mission success and the human and/or the technology are used to achieve success. The first is to assist in the allocation of function between the human and technology in advance of detailed human factors assessment of systems design. Here it is used to predict likely human performance and combine it with likely technology performance to predict mission success. The second is for safety assessment of detailed or completed systems designs, in order to understand and predict the likely levels of system performance in the event of human error and/or system failures. In both applications, the concern is to predict overall performance in terms of the probability of failure, within a factor of ten.

3. PROCESS

The generalised process for performing a HRA___is described in Figure 1. HRA toos, such as PHRASE 2, are primariliy used for steps 4 and 5. The stages are expanded upon below, using PHRASE 2 as an example tool.

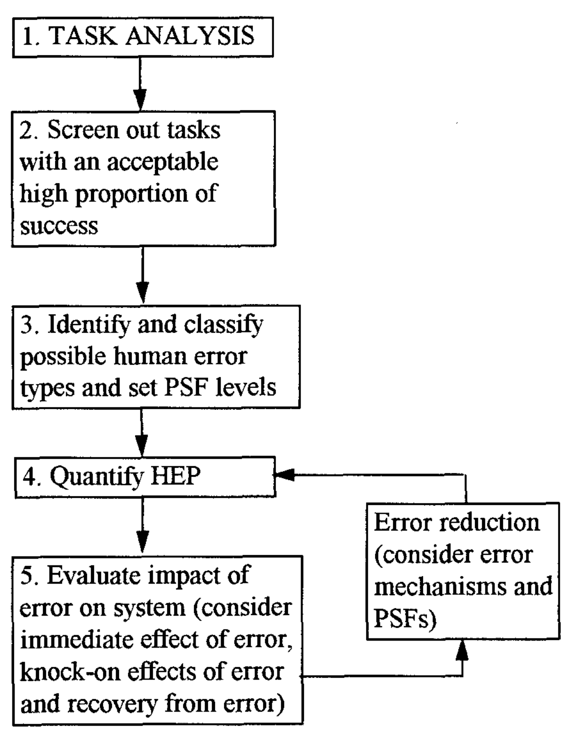


Figure 1: Generalised HRA Process

i. Task analysis.
The first stage is to generate a task analysis and representative scenarios. This is usually performed prior to using a HRA tool. Task analysis is a vital step in order to understand what the user is expected to do. A task analysis performed as a precursor to a

HRA is likely to have a slightly different emphasis than a traditional task analysis, because the analyst will specifically focus on analysing the tasks and task conditions which lead to human error. In PHRASE 2 only task names and task descriptions are recorded - the task analysis and scenarios must be recorded elsewhere.

ii. Determine scope of HRA analysis.
At this stage, because time and resources are always limited, the scope of the analysis can be focused on those tasks that are central to mission success, and to those tasks where the consequences of human error are minimal. This is a very important step in the analysis, and should be performed by an expert or team of experts who are familiar with the operational use of the equipment, human factors considerations, and the proposed technology. PHASE provides only basic support for this step in the form of a checklist of questions to help arrive at a set of tasks where task performance of human-machine combination is important for mission effectiveness. Figure 2 below provides an example of one of these checklists.

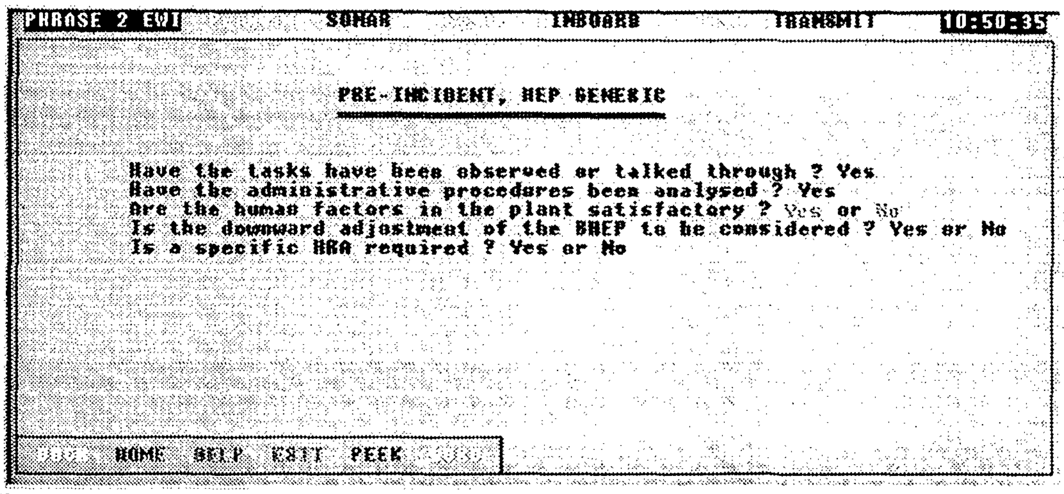
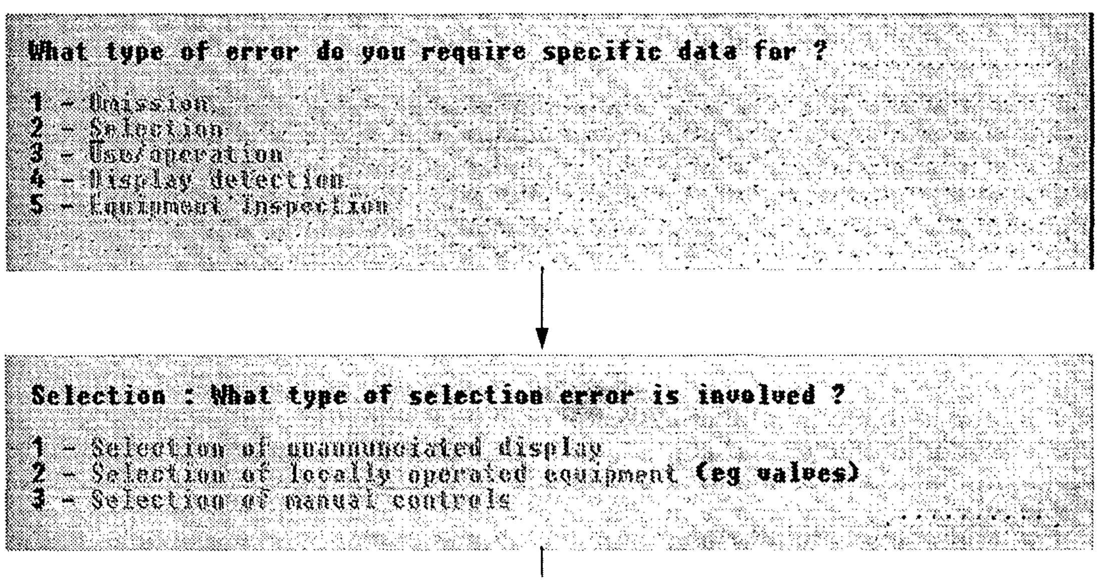


Figure 2: Screen print from PHRASE 2 showing the calculation of a Generic HEP for quickly screening out tasks with an acceptably high probability of success.

iii. Identify and classify possible human error types and set PSF levels.
Possible errors are identified from the task analysis. These errors are then classified. In PHRASE 2 there is a hierarchical taxonomy of error types and the user picks an item from successive lists to classify possible errors. See Figure 3.



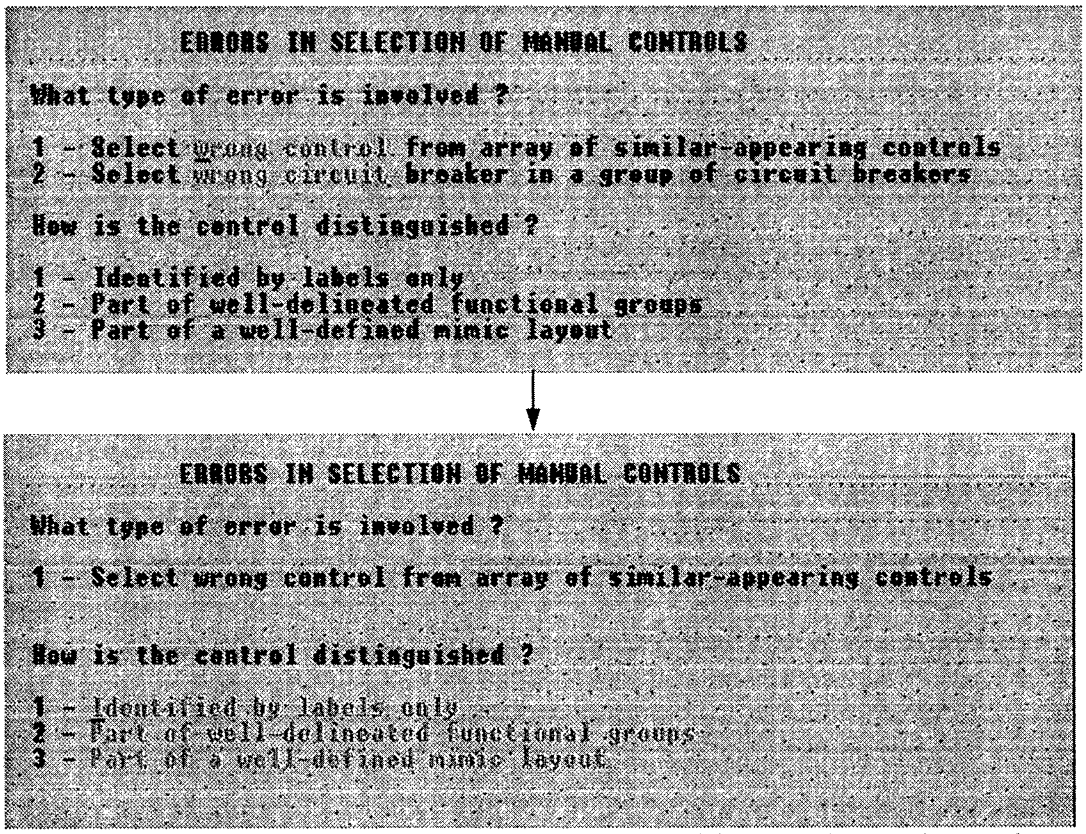


Figure 3: Screen print from PHRASE 2 showing an example of the successive questions used to identify error type (items 2, 3, 1 and 1 were picked from the successive lists).

The Performance Shaping Factors/Performance Influencing Factors (PSFs/PIFs) that affect the triggering of the identified error type are then identified. This is illustrated in Figure 4 below.

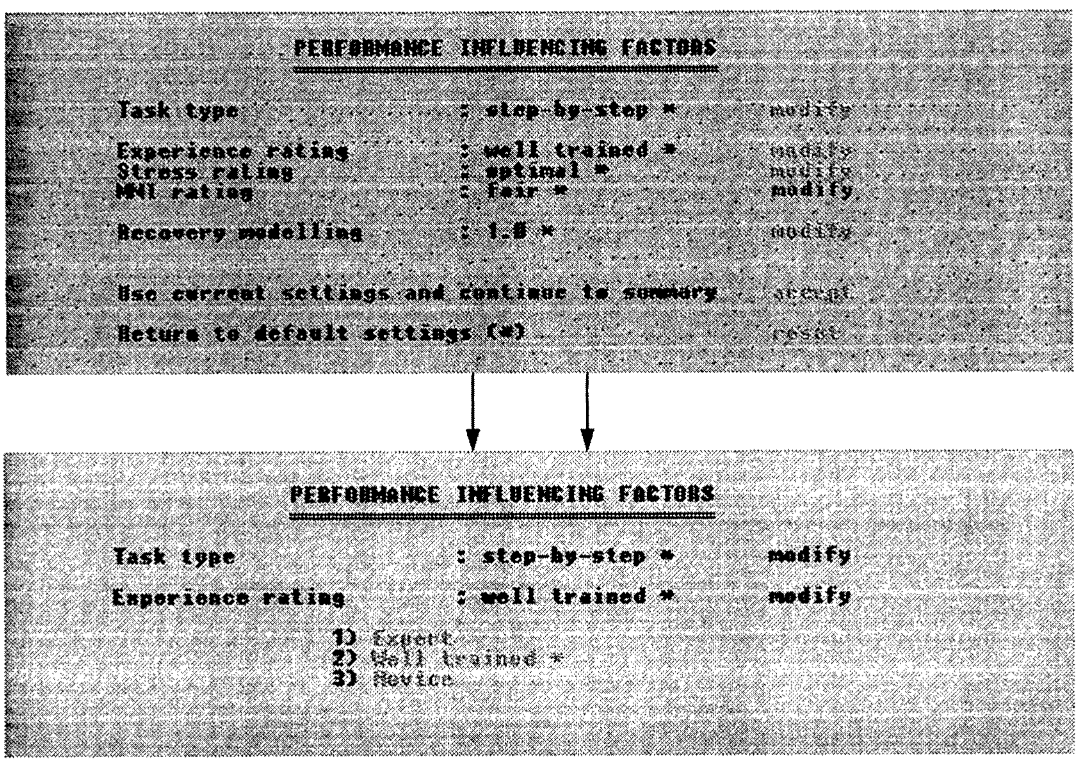


Figure 4: Screen print from PHRASE 2 showing the setting of PSFs/PIFs levels

Dependency can be viewed as a particular type of PSF. The levels of a pre-defined set of PSFs are set in PHRASE 2. See Figure 5 below.

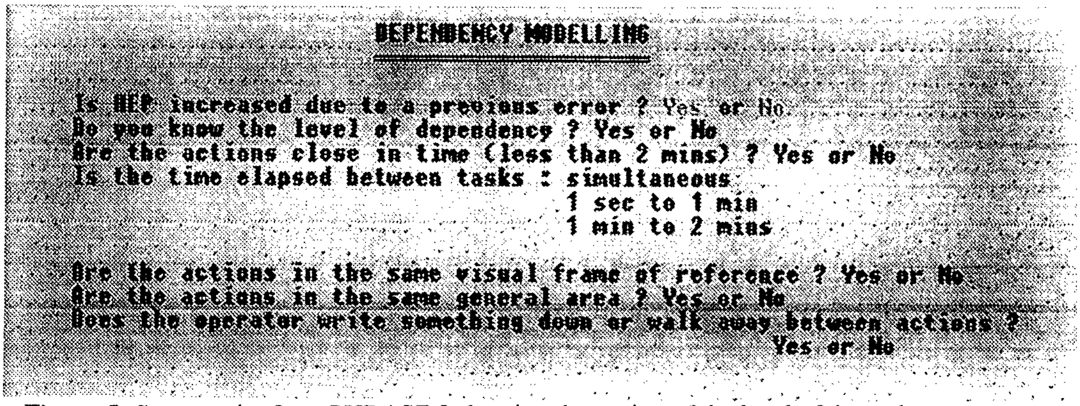


Figure 5: Screen print from PHRASE 2 showing the setting of the level of dependency.

iv. Quantify HEP.
Information about error type and PSFs are then combined to generate an HEP. In PHRASE 2, a basic HEP is extracted from a database for the type of human error defined. This value is then modified according to the prevailing PSFs.

PHRASE 2 classifies all tasks as either pre-diagnostic, diagnostic or post diagnostic. A 'Pre-diagnosis Task' could be either a) a task performed early in the flight or more likely, b) a task performed by ground crew maintenance which leaves the in-flight systems not in their expected state of readiness. A diagnostic task would be diagnosing a system failure in the aircraft or the cognitive performance of a mission oriented task such as identifying a target. PHRASE 2 uses the Human Cognitive Reliability Model here, which assumes that performance on a Diagnosis Task can be predicted from the time available for the task. The 'Post Diagnosis Task' is modelled in the same way as the 'Pre-Incident Task' as described in i-iv above.

v. Evaluate impact of error on system (consider immediate effect of error, knock-on effects of error and recovery from error).
PHRASE 2 allows the calculation of error probabilities for actions required to recover from an error. This can include a diagnostic or a post diagnostic task. However, in order to determine the implications for mission success, either an event tree or a fault tree is commonly used. Event trees start from the basic initiating event and map out the major event sequences leading either to recovery of normal status or to accident conditions. Fault trees tend to look at the combinations of system and operator failures that contribute to the mission failure.

vi. Error reduction (consider error mechanisms and PSFs)
High risk errors can be addressed by redesign. Information on the psychological mechanism behind an error and the sensitivity of the HEP to PSFs levels provide guidance for redesign. PHRASE 2 does not provide information on the psychological mechanisms behind an error, but other HRA tools do (*e.g.*, SHERPA[1]). Other HRA tools also link errors and PSFs to Error Reducing Mechanisms - ERMs (*e.g.*, HEART).

4. Limitations

i. PHRASE 2 and THERP[1] (on which PHRASE 2 is based) have been developed primarily for addressing human error in process control. As such, the use of this specific tool is limited for cockpit design.
ii. The level of system specification must be sufficient to support a detailed task analysis.

It is important to understand that a full HRA requires a detailed task analysis to be in place.

5. Facilities/Resource Requirements

PHRASE 2 runs on a PC and is produced and distributed by Electrowatt Engineering Ltd:

Electrowatt Engineering Ltd.
North Street
Horsham
West Sussex
RH12 1RF
UK
Tel. 44 (0)1403 250131

HEART is an alternative to THERP and is an example of a HRA technique that has been developed to be quick, simple to use and easily

understood. This is achieved by concentrating only on those ergonomic factors which have a large effect on performance. Electrowatt Engineering Ltd. also markets HEART-PC (based on the HEART technique) which is probably better suited to most military context as it is application independent. It is also better suited to dealing with tasks with a strong cognitive component as it focuses on psychological features of the task rather than features of the MMI and environment. a HEART tool.

6. REFERENCES

1 Kirwan, B., A practical guide to Human Reliability Assessment, 1st Edn, Taylor and Francis

Case Study Involving FAIT

Dr. Victor Riley
Honeywell Technology Center—MN65-2600
3660 Technology Dr.
Minneapolis, MN. 55418
U.S.A.

riley@htc.honeywell.com

1. SUMMARY

This document describes in detail the capabilities of Honeywell's Function Allocation Issues and Tradeoffs methodology, its assumptions and philosophy, methods of use, and types and utility of output. This case studies illustrates the process and applicability of FAIT in evaluating the potential human factors issues inherent in a proposed piece of aircraft automation: a new implementation of data link technology.

2. PROBLEM SPACE

One of the most difficult parts of developing a new, complex system is anticipating the full range of human factors issues and possible errors, and designing to prevent them. The Function Allocation Issues and Tradeoffs (FAIT) methodology is intended to assist in this process by making it as systematic and comprehensive as possible. Because it uses a general model of human-machine interaction instead of a system-specific architecture description, it can be applied very early in the concept development stage, before the specific design of the system has been set. This facilitates the definition of requirements that address human factors issues and prevent the types of errors that might otherwise be committed.

The data link system for communications between air traffic controllers and pilots is an example of a system that can benefit from this type of analysis. It is safety critical, so potential human errors must be identified as early as possible and prevented through design solutions as completely as possible. It is also highly complex, with many participants, both human and automated; this complexity implies that there are many possible sources of confusion and error that should be considered. At the time of this analysis, the requirements for data link systems and protocols were being defined by industry committees, and candidate architectures were being developed and user interfaces designed and tested. Application of a FAIT analysis was intended to facilitate and inform this process.

3. DESCRIPTION OF PROCESS

The FAIT methodology is purely analytical. No special software is required. Analysis begins with the identification of what levels of automation are appropriate for the system under consideration, using a taxonomy of automation defined by Riley [1]. The levels of automation determine the form of a general model of human-machine interaction best suited to represent information flow between the operational environment, the automation , and the human operator(s) involved in the process. This model is then used to decompose the system into characteristics of the environment, the machine(s), and the operator. Once these characteristics have been defined, they are placed along both sides of a matrix and potential requirement relationships and real-time interactions are identified between all pairwise combinations of characteristics. In the process of identifying these relationships and interactions, the analyst must construct mental scenarios implied by each combination. This often leads to the identification of potential failures and errors and their possible consequences. These descriptive results are often the most valuable results because they lead to specific requirements.

In addition, the total numbers of interactions along both dimensions of the matrix can be used as an indicator of the importance of each characteristic in the system, in terms of how influential and sensitive each characteristic is in system operation. Characteristics that are both highly influential and highly sensitive can be likely sources of system instability and merit special attention. Finally, symmetrical interactions between characteristics indicate tradeoffs that must be decided during the design process. The results of a FAIT analysis include a description of potential failures and errors and the conditions that may lead to them, measures of influence and sensitivity for each characteristic, and identified tradeoff areas. The descriptions of failures and errors lead directly to the definition of system design requirements. The measures of influence and sensitivity can assist in

project planning, indicating where design resources should be concentrated. Tradeoff areas lead directly to the identification of trade study topics.

The following sections describe this process in more detail.

3.1 Inputs

There are two types of inputs to a FAIT analysis: external and internal. The external input is the general description of the system to be analyzed. In the case of data link, the system consists of the communications equipment used (including the controllers' and pilots' communication devices), the messages being sent, the sensors through which messages are received, the workstations of the human operators (but at a very high level: knowledge that the workstation has displays and controls and the functions that they generally perform is adequate), and the types of procedures likely to be followed. Note, however, that since the methodology is intended for very early front end development, this description can be very high level and general. Knowledge of the system architecture, the specific types of control and display devices and their exact functions, the types of sensors to be used, and the specific procedures is not needed. In the case of data link, merely knowing that the system is intended to replace much of the voice communications between air traffic control (ATC) and aircraft, that there may be transmission delays, that messages are likely to be presented on a visual display but may be presented aurally, that the flight guidance parameters contained in messages may be "gated" into the flight guidance system so the pilot doesn't have to manually enter the data, that messages are only presented to the specific aircraft to which they are sent, and that one pilot is likely to have communications responsibility while the other pilot flies the aircraft is sufficient for a FAIT analysis. Indeed, if more specific information were required, analysis would have to wait until the system was further along in its design, and the results would have less value because corrections are more expensive to make the more mature the system definition is.

The internal inputs are the information that the FAIT process itself brings to the analysis. These include a taxonomy of automation, a general model of human-machine systems, and predefined, reduced forms of this model that correspond to all the possible combinations of automation levels in the taxonomy. The taxonomy, shown in Figure 1, provides several levels of automation "autonomy", which refers to the level of permission the automation has to manipulate information and take action, and "intelligence", which refers to the type of information the automation can use.

LEVELS OF INTELLIGENCE

LEVELS OF AUTONOMY	raw data	procedural	context responsive	personalized	inferred intent responsive	operator state responsive	operator predictive
none							
information fuser							
simple aid							
advisor							
interactive advisor							
adaptive advisor							
servant							
assistant							
associate							
partner							
supervisor							
autonomous							

Figure 1: The taxonomy of automation

Theoretically, any human-machine system can be represented as a combination of a level of autonomy and a level of intelligence. For example, a simple radar display, which merely indicates the presence of a returned signal from some remote object, is at the "none" level of autonomy and the "raw data" level of intelligence. If that signal is processed to include data specific to the targets being represented, such as an ATC display that shows flight information in data blocks attached to enhanced returns, the level of autonomy rises to "information fuser" due to the added information and the level of intelligence rises to "procedural" to reflect the need for additional processing.

To identify the levels of autonomy and intelligence that are appropriate for the system under consideration, the analyst provides answers to two lists of questions about the system's capabilities. For each question, a "no" answer indicates that the current level is not an appropriate descriptor and the analyst should continue down the list. A "yes"

answer indicates that the appropriate level has been found. The questions are as follows.

3.1.1 Autonomy Questions

These questions are used to determine the system's level of autonomy. The analyst starts with Question 0 and follows the instructions.

0. Does the machine perform any control actions?

If the machine can only manipulate information, answer "no" to this question. If it performs some type of actions, answer "yes". Examples of machines that do not perform control actions are radar, voice radio, display devices, and caution and warning systems. Examples of systems that can perform control actions are the autopilot, flight management system, and system controllers that can reconfigure systems automatically.

If the answer to this question is "no", go to question 7. If the answer is "yes", then continue with question 1.

1. Does the machine act without informing or interacting with the operator?

If the answer to this question is always "yes", the level of autonomy is Autonomous. Because the system is autonomous, there is no human operator to consider, and no human factors issues are relevant. However, as part of a larger system, there may indeed be human factors issues. For example, if the aircraft performed all fuel management automatically and the crew had no displays of fuel state and no indication of what the automation was doing with the fuel, the relevant human factors issues would arise at the level of functions where the operators are involved and fuel is a related concern, such as navigation, and not at the level of fuel management. If this is the case, try to determine what larger system the subsystem you have in mind is part of and start over, considering the original system as a subsystem of the larger system. This will incorporate the human operator into the picture, and the autonomy of the original subsystem will be one of the characteristics you would include for it.

If the answer to this question is "no", go on to question 2.

2. Does the machine have more authority over the operator than the operator has over the machine?

If the machine can override the human operator, but the operator cannot override the machine, answer "yes" to this question. Select the Template for Supervisor in Figure 3 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 3 of this list.

3. Do the machine and the operator have roughly equal authority over the other?

If the machine can override the operator sometimes and the operator can override the machine sometimes, answer "yes" to this question. Select the Template for Partner in Figure 4 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 4 of this list.

4. Can the machine take over operator tasks automatically?

This refers to a capability of the machine to recognize when the operator needs assistance, is about to make an error, or is doing a task poorly, and to automatically take over the task without being specifically directed to by the operator. In this case, the operator can override the system if necessary, but the system has the authority to take over operator tasks without being asked to. If the answer to this question is "yes", select the Template for Associate in Figure 5 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 5 of this list.

5. Can the machine take over operator tasks with standing permission or consent?

This refers to automation that can perform selected tasks when the operator directs it to do so, and continues to perform the task until the operator takes the task back or directs it not to do the task. An example of this level of automation would be LNAV, which performs the manoeuvres necessary at each waypoint to stay on the planned track. If the answer to this question is "yes", select the Template for Assistant in Figure 6 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 6 of this list.

6. Can the machine take over operator tasks when the operator explicitly hands them off to the machine?

This refers to automation that can perform selected tasks on a case by case basis. The operator must direct it to perform the task each time. An example of this level of automation would be the heading control on the glareshield, which performs one manoeuvre to orient the aircraft to the selected

heading. To change the heading again using the same function, the operator would have to enter a new heading in. If the answer to this question is "yes", select the Template for Servant in Figure 7 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 7 of this list.

7. Can the machine manage the operator's displays autonomously?

This refers to automation that does not perform any control functions but can completely manage the presentation of information to the pilot, such as determining what information should be presented, what format it should be presented in, and how it should be presented. If the answer to this question is "yes", select the Template for Adaptive Advisor in Figure 8 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 8 of this list.

8. Can the machine initiate interactions with the operator?

This refers to automation that can make recommendations to the operator without being explicitly asked for them and request information from the operator, but it does not have the authority to filter information the way the Adaptive Advisor does. If the answer to this question is "yes", select the Template for Interactive Advisor in Figure 9 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 9 of this list.

9. Can the machine provide recommendations or advice?

This refers to automation that can provide recommendations to the operator when the operator asks for them and when the automation recognizes that the circumstances are right for making a recommendation. If the answer to this question is "yes", select the Template for Advisor in Figure 10 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 10 of this list.

10. Does the machine perform any decision making functions?

This refers to automation that assists the operator by providing decisions on a case-by-case basis. An example would be an automatic target recognizer that attempts to categorize radar returns as belonging to targets. If the answer to this question is "yes", select the Template for Simple Aid in Figure 11 and go on to the Intelligence Questions in Section 3.1.2. Otherwise, continue with Question 11 of this list.

11. Can the machine integrate information and construct displays?

This refers to automation that can collect information and put it in the best format for presentation to the operator. If the answer to this question is "yes", select the Template for Information Fuser in Figure 12 and go on to the Intelligence Questions in section 3.1.2. Otherwise, continue with Question 12 of this list.

12. If the answers to all the above are "no", then select the template for None in Figure 13 and go to the Intelligence Questions in Section 3.1.2.

3.1.2 Intelligence Questions

These questions are used to determine the system's level of intelligence based on the information it can use. The analyst starts with Question 1 and follows the instructions.

1. Can the machine predict the operator's behaviour?

This refers to automation that can use information about the operator's physical state, infer the operator's intent, and anticipate the next actions to be made by the operator. If the answer to this question is "yes", select the Template for Operator Predictive in Figure 14 and go on to "identify characteristics" in Section 3.2. Otherwise, continue with Question 2 of this list.

2. Can the machine monitor the operator's physical state?

This refers to automation that can determine when the operator is unconscious, fatigued, or otherwise physically impaired. If the answer to this question is "yes", select the Template for Operator State Responsive in Figure 15 and go on to "identify characteristics" in Section 3.2. Otherwise, continue with Question 3.2 of this list.

3. Can the machine infer the operator's intent?

This refers to automation that can dynamically infer the operator's intentions and assist the operator to carry them out. If the answer to this question is "yes", select the Template for Operator Intent Responsive in Figure 16 and go on to "identify characteristics" in Section 3.2. Otherwise, continue with Question 4 of this list.

4. Does the machine use embedded models of the operator?
This refers to automation that can be "personalized" to perform or present information the way a particular operator wants it. In this way, in can be thought of as containing a model of the operator that is static and therefore does not change with time or circumstances the way the Operator Intent Responsive model would. If the answer to this question is "yes", select the Template for Personalized in Figure 17 and go on to "identify characteristics" in Section 3.2. Otherwise, continue with Question 5 of this list.

5. Is the machine's output contingent on the state of the situation?

This refers to automation whose behaviours or responses can change based on the situation. An example would be a flight management system which performs different actions depending on the aircraft's location along the flight plan. If the answer to this question is "yes", select the Template for Context Responsive in Figure 18 and go on to "identify characteristics" in Section 3.2. Otherwise, continue with Question 6 of this list.

6. Does the machine operate according to a fixed set of procedures without regard to the situation?

This refers to automation that responds only to internal settings and programming. An example would be a weather radar display that develops a visual code for the display based on the amount of precipitation sensed. If the answer to this question is "yes", select the Template for Procedural in Figure 19 and go on to "identify characteristics" in Section 3.2. Otherwise, continue with Question 7 of this list. 7. If the answers to all the above are "no", then select the template for Raw Data in Figure 20 and go to the Step Two, Identify Characteristics in Section 3.2.

3.1.3 Using the Questions to Select a Model

For the data link system under consideration, the analyst would answer "yes" to the autonomy question relating to the "servant" level and to the intelligence question relating to the "context responsive" level. This is because the system only takes action when the pilot consents, on a case by case basis (when the controller sends up a data link message with embedded guidance parameters, these parameters are only sent to the autopilot for execution when the pilot accepts the message and explicitly send the parameters on), and because the behaviour of the system depends on the current situation (for example, the clearances received by one aircraft depend on the locations of other aircraft, weather, restricted areas, and other constraints on the flight path from the present position).

The other element of the internal inputs, the general model, is shown in Figure 2 on the following page.

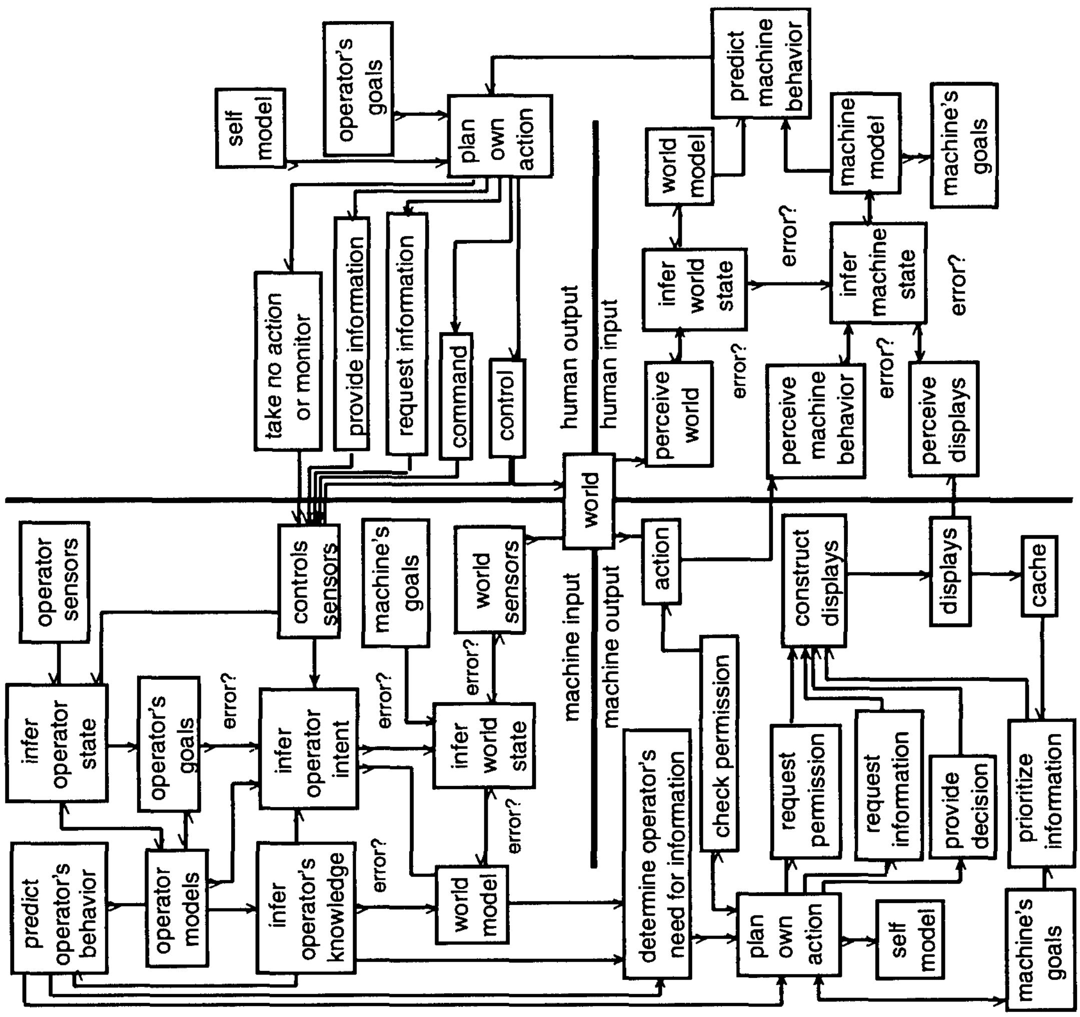


Figure 2: The general model of human-machine systems

The key to using this model to represent the information in a specific human-machine system is that each combination of a level of intelligence and a level of autonomy corresponds to a predefined, reduced form of the general model. In the case of data link, which is at the servant level of autonomy and the context responsive level of intelligence, that reduced form is as shown in Figure 3.

Having identified the model appropriate for the system under consideration, the analyst must now start to apply the technique.

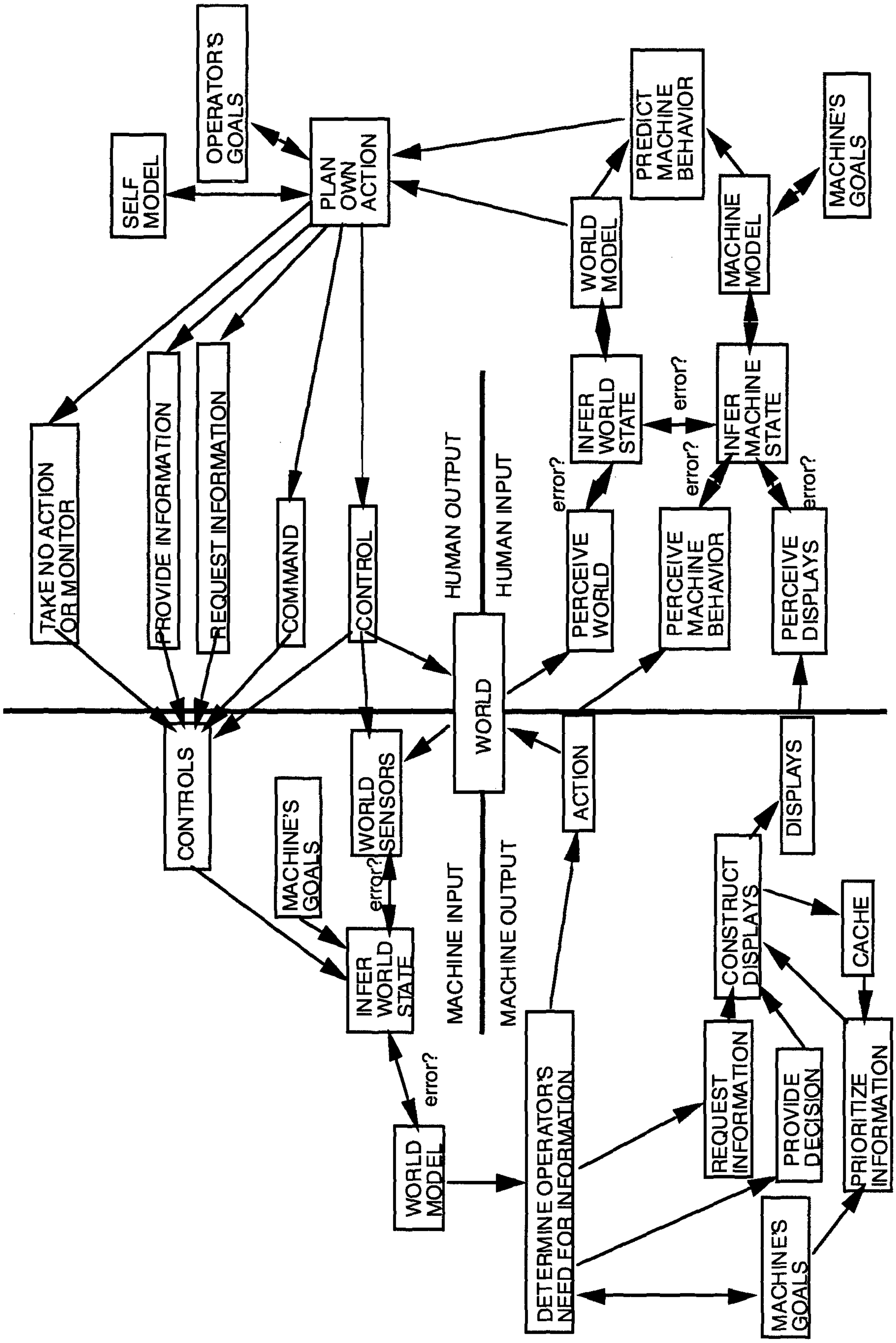


Figure 3: The reduced model form for data link.

3.2 Walkthrough

With the above model depiction in hand, the analyst begins the process of using the model to decompose the system into its component characteristics. There are several benefits to using this model formulation for this process instead of a more conventional system architecture diagram. First, because a system-specific architecture description is not needed, the process can be applied before one is developed, and the eventual architecture can take account of the results of the analysis. Second, a traditional architecture diagram leaves out the operational environment and human operator. Because issues and errors can arise from interactions between all three parties (environment, machine, and operator), the environment and operator must be included in the analysis. And third, because the focus of the analysis is on human factors issues, the representation of the human operator's perceptual, decision making, and response processes must be consistent with the representation of the machine's sensing, information processing, and output processes. This enables a comprehensive analysis of human-machine interactions.

The decomposition of the model into characteristics is performed by considering each box, or node, in the model and identifying everything that might be important about that node. For example, aspects of the operational environment (or "World") that might be important to data link operations include the current position and vector of the aircraft being modelled, the positions and vectors of other aircraft, the weather, visibility, locations of restricted areas, and so forth. Aspects of the machine's Sensors that might be important include the delay times imposed on messages sent through the system. In the case of a Mode S data link system, this delay can be up to four seconds for a message with a size that allows its transmission in a single pass of the Mode S radar, longer for larger messages. Aspects of the pilot's Perceive World node that might be important include how much time the pilot has to look out the windscreen for other aircraft (head up time) and access to messages sent to other aircraft, a feature of the current voice communications. The pilot's level of situation awareness is represented as a characteristic in the pilot's World Model node, and workload is represented in the Plan Own Action node. Display reading delays and accuracy are represented in the pilot's Perceive Displays node, and control input delays and accuracy are represented in the pilot's Command and Control nodes. Figure 4 shows how some of the characteristics map into the nodes in the model.

A full list of the characteristics identified for the data link analysis follows:

The World node in the model contains those characteristics of the operating environment relevant to data link that may be of interest for human factors issues. These characteristics may include the positions and velocities of aircraft in the immediate area and their types and quantity, weather, and the availability of clearances. Aircraft positions and velocities may be important for sector capacity and error recovery considerations. Aircraft type may be important because of mixed environment considerations, in which some aircraft will have data link and others not. Weather may be important due to the constraints it may impose on operations, the effect it may have on pilot visibility, and other factors. And clearance availability may be important due to the operational constraints acting on the controller.

The World Sensors node contains sources of information coming into the aircraft systems that may have an effect on data link operations. These include radar (traffic and weather), satellite data, automated information sources, and data from other pilots and aircraft. Characteristics of interest include the rates at which such information is updated, what data are available, and whether the information arrives in data or voice form.

The Infer World State node on the Machine side of the model represents the air traffic controller's or company dispatcher's process of determining the current state of traffic and conditions. Characteristics of interest include the accuracy of the information available, the processing delay incurred as the controller or dispatcher mentally sorts out and interprets events and information, controller or dispatcher workload, and controller or dispatcher coordination with the aircraft under his or her responsibility. The Machine Goals node represents the influence on this step exerted by the controller's and company's goals, in this case the desired aircraft behaviour.

The World Model node on the Machine side of the model represents the controller's and dispatcher's mental model of the situation, or the controller's and dispatcher's situation awareness.

The "Determine Pilot's Need for Information" node represents the controller's and dispatcher's process of deciding what information to send on to the pilot through data link.

The "Plan Own Action" node on the Machine side of the model represents the processing performed by the system to implement the information received as a flight control action. Since this is strictly a gate in the case of data link, there are no characteristics of interest here.

The "Request Information" and "Provide Decision" nodes reflect the authority of the system to initiate certain types of transactions with the pilot. Characteristics of interest relevant to all of these are the priorities and delays incurred during the transmission of information from the data sources to the aircraft, such as the four second Mode S transponder delay.

The "Action" node represents the control action

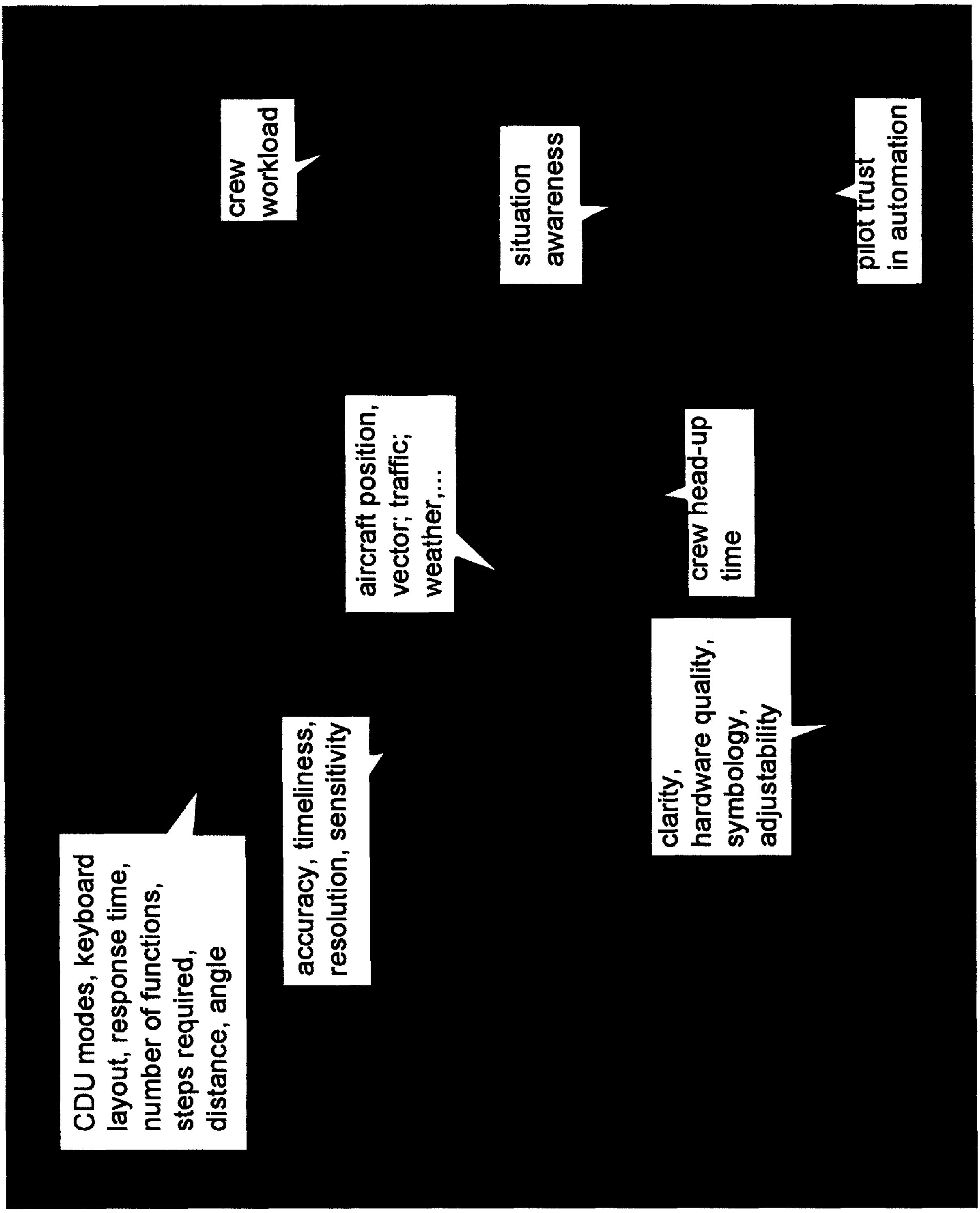


Figure 4: Some characteristics of data link

taken by the FMS after information or a command has been gated into it from the data link system. This action changes the situation in some way, hence the output back to the World node.

The "Construct Displays" node represents the process of determining which message to display and what format to use. A characteristics of interest here is the availability of display space, or the potential interference or competition between functions for shared display space.

The "Prioritize Information" node represents the mechanism by which the Construct Displays function determines the order in which to queue waiting information. Characteristics of interest here are the particular prioritization scheme adopted for data link and how well the eventual priority assignments fit with the pilot's own priorities.

The "Cache" is a temporary storage area for information waiting in the display queue. A characteristic of interest here is the length of time a piece of information waits to be read.

The "Displays" node represents the physical display devices by which data link information is displayed to the pilots. A characteristic of interest here is display clarity, which refers to both the physical readability of the display and the perceptual and conceptual clarity of the human-computer interface design.

The remaining nodes in the model represent the perceptual, mental processing, decision, and response functions of the pilot interacting with the data link system.

The "Perceive World" node represents the pilot's ability to derive information directly from the world by looking through the windscreen, monitoring radio transmission, and using the Traffic Collision Avoidance System (TCAS) display. Characteristics of interest include the pilot's field of view, external visibility, how much head up time the pilot has, and the "party line" open voice radio channel. This is an important consideration for data link because pilots may lose the awareness of communications between other aircraft and ATC when they receive only messages intended for them.

The "Perceive Machine Behaviour" node represents the pilot's ability to detect changes in aircraft behaviour resulting from a control action. Specifically, the pilot may detect a manoeuvre performed by the autopilot after a new command has been entered into the flight guidance system or after the FMS initiates a manoeuvre in response to the flight plan.

The "Perceive Displays" node represents the pilot's display reading. Characteristics of interest include the pilot's level of skill in interpreting the display, the delay associated with the pilot's finishing other activities before reading the display, reading errors, and the workload demand imposed by the display which may arise both from the complexity of the human computer interface design and from the physical placement and visual characteristics of the display hardware.

The "Infer World State" node on the Pilot side of the model represents the construction of the pilot's mental model of the situation. The characteristics of interest here include the accuracy of the information, the delay incurred by the pilot's process of integrating the information from several sources and interpreting it, and coordination between the crew as they confirm each other's understanding of the situation and share opinions about it.

The "World Node" on the pilot's side of the model represents the pilot's mental model of the situation and data sources. The characteristics of interest here are the accuracy of that model, or the pilot's level of situation awareness, the amount of risk the pilot attributes to the situation, the degree to which the pilot trusts the data received, and the pilot's assessment of the urgency of a communication.

The "Infer Machine State" node refers to the pilot's inference of how trustworthy the data link system is.

The "Machine Model" node represents the pilot's ability to anticipate requests and clearances.

The "Plan Own Action" node on the pilot side of the model represents the pilot's decision making. Characteristics of interest here include the pilot's workload and the time it takes to arrive at a decision.

The "Pilot's Goals" node represents the pilot's preferences for particular clearances and the pilot's prioritization of the constraints that determine what actions and responses are possible.

The "Self Model" node on the pilot's side represents the pilot's opinion of his or her own abilities. The characteristic of interest here is the pilot's level of self confidence, or confidence in their own ability to handle a given situation appropriately.

The pilot output nodes, particularly the "Command" and "Control" nodes, represent the types of responses the pilot may make. Control responses represent direct inputs to the flight control system, and command responses represent information inputs to the flight guidance system or to ATC, including through data link. Characteristics of interest here include the delay associated with making a response, errors, workload demands imposed by the human computer interface design and physical placement and characteristics of the control devices, the ease of using the devices, and interference due to competing demands for the same devices, as may arise if data link is shared with other functions on the FMS Control Display Unit.

The "Controls Sensors" node represents the interface between the pilot's control devices and the aircraft's information management systems.

It should be noted that some of these characteristics (such as workload) are generic to most human-machine systems while others (such as the voice radio "party line") are specific to the data link analysis. Again, the purpose of the automation taxonomy and information flow model is to guide the analyst in identifying characteristics as systematically and comprehensively as possible; for the most part, the process does not provide the characteristics itself.

Having identified the characteristics to be considered in the analysis, the analyst then lists the characteristics along both side of a matrix. The structure of the matrix is shown in Figure 5.

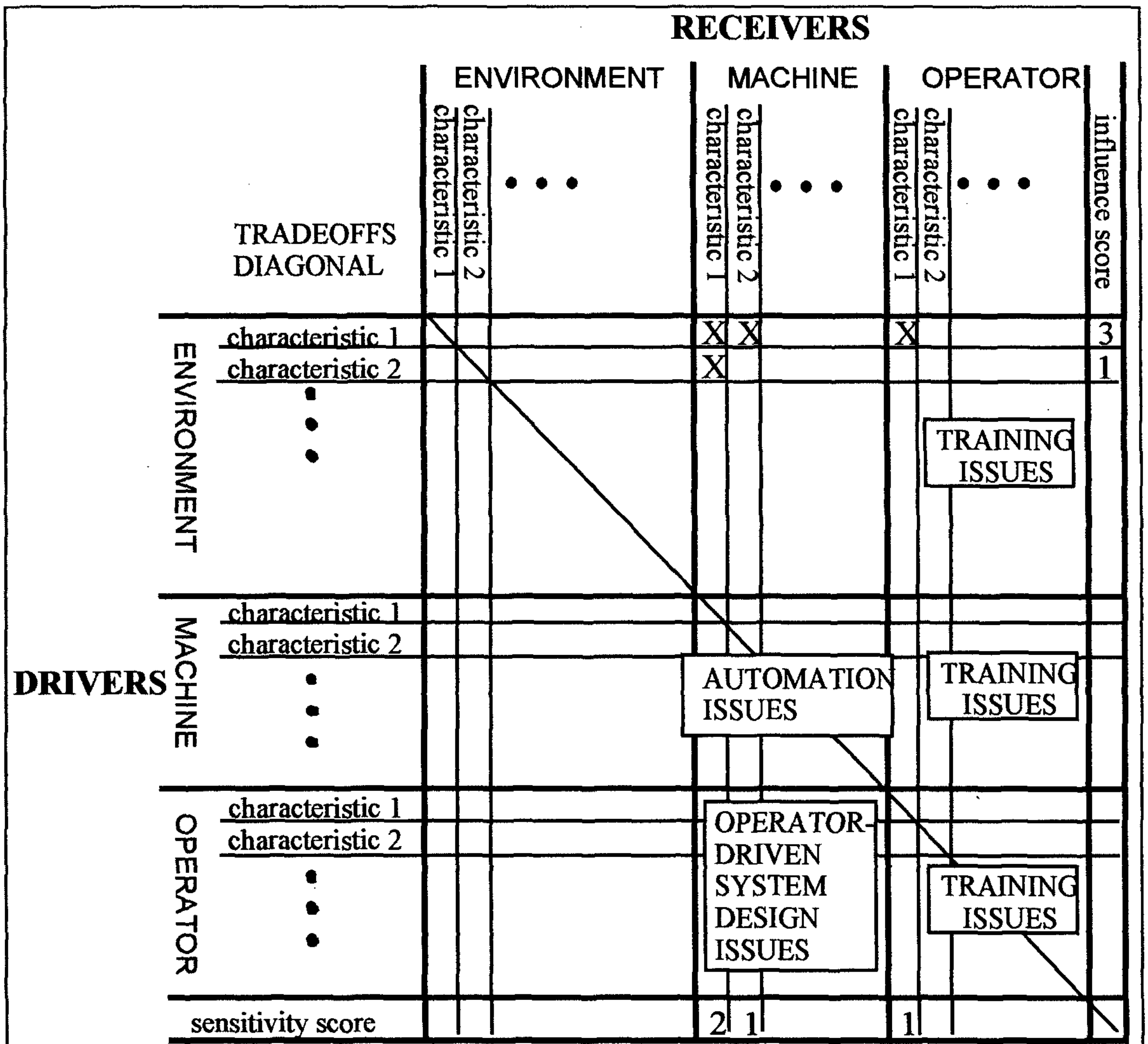


Figure 5: The structure of a FAIT analysis

The purpose of this matrix is to enable the analyst to comprehensively consider the potential interactions and relationships between all pairwise combinations of characteristics. Each possible pair is considered twice (except for those along the diagonal) because each characteristic must be considered as both the driver of an interaction or requirement and as the receiver of one. To illustrate the difference, the analyst must ask two questions for each cell of the matrix: first, can the row characteristic (the driver) affect the column characteristic (the receiver) during real time operation of the system; and second, can the row characteristic place a requirement on the column characteristic. To answer these questions, the analyst must often construct mental scenarios in which the answer to either question might be "yes". If such a scenario is found, the analyst enters a mark into the associated matrix cell and documents the scenario and any attendant failures, errors, and issues.

3.3 Outputs

These scenario descriptions and their associated requirements and issues usually constitute the most valuable and directly usable results of a FAIT analysis. Having such a broad and deep view into all the possible interactions between and within the world, the machine, and the human operator affords a very extensive identification of relevant issues. Some sample issues for the data link analysis follow. Because the topic of situation awareness is of great interest in the human factors community, and because anticipating the effects of a system concept on the situation awareness of the human operators has been very inexact, situation awareness issues resulting from the data link analysis [2] are presented below to illustrate the ability of FAIT to help the analyst grapple with a very abstract topic area. For each issue description, the pair of characteristics that gave rise to the issue are noted with the driving characteristics first and the receiving characteristic second.

3.3.1 Information Update Rate -

Situation Information Accuracy: Acknowledgements, queries, and information received from flight crews constitute several types of information that come into the controller's or dispatcher's task. To the extent that flight crews differentially delay their acknowledgements or that a controller delays reading information sent by flight crews, the controller may be required to integrate information from over a relatively large range of time. This is in contrast to present practice in which the controller's or dispatcher's verbal dialogue represents a single time referent. In other words, where controllers or dispatchers currently deal with the present and future, they will also have to deal with the past using data link.

3.3.2 Voice vs. Data -

Crew Coordination: Voice communications are received by both pilots, promoting a common level of situation awareness between pilots. It may be possible for one pilot to read, acknowledge, and even enter data link information into the FMS without the other pilot's knowledge. This potential should be considered and addressed in the design of the data link human-computer interface and flight deck procedures for using data link. Crew Resource Management (CRM) strategies should be implemented to ensure that both pilots share a common view of the situation and continue to cross check each other for errors.

3.3.3 Voice vs. Data - Recognize Urgency Level:

Voice communications permit the flight crew to receive implicit information, such as inferring the urgency level of a request by the controller's tone of voice. Data link may not provide the wealth of implicit communication that voice does.

3.3.4 Voice vs. Data - Pilot Trust in Data:

It is commonly thought that people attribute greater reliability to numbers on computer screens than they do to written or spoken numbers. The tendency to over-rely on machines has prompted the phrase "garbage in, garbage out". It is possible that pilots will also place more trust in information appearing on data link screens than transmitted over voice channels and therefore not perform careful or extensive error checking on such data, leading to more ready acceptance of erroneous data. It is also possible that the easy gating of data into the FMS may result in more flight path deviations due to controller errors that were not caught by the flight crew.

3.3.5 Controller Workload - Voice vs. Data:

There may be a tendency for a controller to revert to all-voice communications under high workload or high risk periods. This may arise for several reasons: the need for high situation awareness and a single time referent for all communications (that is, a single line of dialogue rather than many messages

subject to various amounts of waiting time); the need for immediate feedback from aircraft; the greater difficulty associated with managing a mixture of voice and data communications; and, at least initially, force of habit. How will such a tendency affect communications, pilot expectations, and ATC/aircraft coordination during such periods, and particularly during the transitions as the controller begins to rely more heavily on voice at the start of the period and less so toward the end? How will a smooth transition back to the desired mixture of data and voice communications be accomplished when the high workload or risk period is alleviated?

3.3.6 Controller Workload - Action Errors:

Controller errors may be induced by high controller workload. The conditions that create high controller workload may also be creating high flight crew workload. This may result in a situation where ATC is more likely to produce an error and the flight crew is more likely to gate the error into the FMS because they don't have the time to check the information adequately.

3.3.7 Controller Workload - Risk Assessment:

Pilots often adjust their methods of interacting with ATC based on their assessment of the controller's workload and stress levels (such as deferring low priority requests). The lack of information about controller or sector workload may prevent pilots from facilitating information flow this way and lead to greater demands on the controller. Some pilot indicator of sector workload may be useful.

3.3.8 Gate Availability - Pilot Situation Awareness:

Because being able to gate data linked data directly into the FMS removes the need for the pilots to directly enter the data themselves, it may be possible for a crew to almost automatically gate data in out of habit without thoroughly checking it and understanding it. This may facilitate flight crew laziness about the data and eventually reduce their situation awareness.

3.3.9 Fit With Pilot's Priorities - Action Error:

If a crew expects a particular clearance, receives a different one, but fails to notice the difference and gates the data in question into the FMS, and the aircraft may behave differently than the flight crew expects. However, this may be a benefit; the aircraft will behave correctly even though the flight crew is in error.

3.3.10 Party Line - Request Anticipation:

Data link may remove one of the sources of information, ATC transactions with leading aircraft, that enable flight crews to anticipate clearances and requests.

3.3.11 Party Line - Workload:

Pilot workload may drop partly as a consequence of reduced opportunities to gain situation awareness, but pilots may incur greater workload attempting to make up the situation awareness deficit from other sources.

3.3.12 Crew Coordination - Controller Situation Awareness:

While voice communications are audible to both pilots, it is possible that one pilot may read and acknowledge a data link transmission without verifying that the other pilot is fully aware of the transmission and agrees with it. This lack of crew coordination may result in a situation where the controller believes the message will be complied with but the pilot flying does not comply with it, resulting in behaviour unexpected to the controller.

3.3.13 Crew Coordination - Action Errors:

It is also possible that the pilot not flying may gate data linked data into the FMS without the pilot flying positively confirming and being aware of the action. This may result in aircraft behaviour unexpected to the pilot flying.

3.3.14 Request Anticipation - Display Reading Error:

A flight crew with a high degree of confidence in their expectation of a clearance or request may let the clearance or request message wait longer than normal, causing a recovery action if the clearance or request is not as expected.

3.4 Using Matrix Results

This represents a very small fraction of the total number of issues resulting from the FAIT analysis. In addition to situation awareness, these issues included display and control design and placement, display formatting, transmission delays and errors, crew coordination, workload, alerting, procedures,

automation levels and functioning, pilot head up time, and many others.

In addition to the scenario and issue descriptions, the requirements responses lead directly to requirement areas, or topics for which specific requirements must be written. The marginal totals along both axes of the matrix indicate how influential each characteristic is in the operation of the system and how sensitive each characteristic is to influence from other characteristics. Characteristics that are both highly influential and highly sensitive are likely sources of instability and should be given careful treatment in the design. In this way, the FAIT methodology can help the system developer allocate design resources to the most important problem areas.

Finally, symmetry around the negative diagonal of the matrix indicates where trade studies should be performed. A typical example of such symmetry is that display quality can place a requirement on operator reading ability, but limitations in operator reading ability can place requirements on display quality. This indicates that the developer needs to trade off a short term investment in system quality against a long term cost in personnel selection and training.

For data link, the results of the analysis were used to drive several industry documents detailing data link issues, recommending a research agenda to tackle the issues systematically, and setting forth human factors requirements for data link systems. The research agenda document was submitted by the Air Transport Association to the FAA for inclusion in the National Plan for Human Factors and was used by the FAA flight deck research office to guide data link research funding. The requirements document was submitted to the RTCA working group developing the Minimum Operational Standards for data link, and the human factors requirements were either included within the body of the document or cited as additional requirements.

4. FACILITY/RESOURCE REQUIREMENTS

The FAIT analysis process does not require any computer equipment, although the analyst may find that the use of a spreadsheet greatly facilitates the construction and manipulation of the matrix. The primary requirements for the analyst are human factors expertise, so the analyst can accurately identify and characterize human factors issues, and general knowledge of the system being considered.

5. REFERENCES

1. Riley, V. "A General Model of Mixed-Initiative Human-Machine Systems." *Proceedings of the Human Factors Society Annual Meeting*, Denver, CO. 1989.
2. Riley, V. (1992). "Human Factors Issues of Data Link: Application of a Systems Analysis." In *Proceedings of the Conference of the Society of Aeronautic Engineers, Aerotech 92*, Anaheim, CA. 1992.

Worked Example Of The Oracle Target Aquisition Model

Dr P Emmerson
Sowerby Research Centre
British Aerospace Defence Ltd
PO Box 5
Filton, Bristol
UK

1.PROBLEM SPACE

A designer is asked to provide a human operator with optimised values for the gain on an electro-optical sensor system in a land fighting vehicle. The gain (or 'temperature window') is known to affect target aquisition, and the designer decides to issues guidelines for the optimum gain for specific situations, based on predictions from a human visual target aquisition model. The chosen model, ORACLE, predicts target aquisition performance under a wide range of conditions, and can include performance with a variety of sensors. For this example, the thermal imaging model is used, in which a single parameter (gain) is iterated over a realistic range for the TI, for a single scenario (a given target and environmental conditions). A complete solution to the designers requirement would involve iterations over other variables (for example different atmospheric visibilities), but all such iterations would follow the procedure outlined below). It is to be noted that this example has been chosen to show the potentially wide range of input parameters that can be used.

2. DESCRIPTION OF THE PROCESS

2.1 Inputs.

2.1.1 Data description

The information below gives a brief description of the inputs required. It should be noted that the model contains default data that will be applicable to many situations. The TI data presented are completely generic, and do not relate to any particular system. In the present example, the designer would be required to have available a fairly complete technical specification of the equipment proposed, but would not need to be concerned with visual or environmental inputs. The inputs that would probably be required are indicated by *. The other variables are documented to show the factors that the model takes into account.

INPUT : Fixation or Glimpse Time
Units : seconds

Fixation refers in this instance to the time during a visual search pattern during which the observer foveates (or inspects) one particular position in the visual scene. The search task will consist of a sequence of such foveations punctuated by rapid eye movements between fixations.

INPUT : Maximum Number of Glimpses for Search
Units : dimensionless

This variable specifies a maximum possible search time according to the relationship :
Maximum Search Time = Maximum No. Glimpses * Glimpse Time

INPUT : Viewing characteristics
Units : dimensionless

The model is configured to run for either monocular or binocular viewing. Binocular viewing is normally associated with a higher performance level than monocular viewing.

INPUT : Confidence Level
Units : dimensionless

The criterion the observer uses to respond to a visual stimulus depends on whether the decision is free or forced. Laboratory studies to establish threshold contrasts often use techniques forcing the observer to respond. Absolute forced choice is represented by a value of 1.0. Free choice where an observer is not in a constrained experimental situation has been experimentally calibrated to be 2.8 times worse and this value of 2.8 is used to represent free choice tasks such as search.

INPUT : Fractional Perimeter
Units : dimensionless

The fractional perimeter is defined as the fraction of the perimeter of a target which it is required be resolvable in order for the observer to successfully accomplish the visual task. The following are the regularly used values which can be related to specific tasks :

Fractional Perimeter = 1 for detection of luminance differences

= 0.45 for tank/bush discrimination
= 0.35 for low clutter scene search
= 0.25 for high clutter scene search
= 0.08 for target identification

INPUT : Search Field Angle
Units : degrees

The model calculates visual performance in 1 degree increments over the full extent of the field of view. The variable therefore sets an effective limit to how far visual performance is calculated into the peripheral visual field and sets a limit to the area for search calculation.

INPUT : Start Range
Units : metres

This variable specifies the target start range in metres. The model may be configured for a target with closing range by specifying a non-zero target velocity. In this case the start range corresponds to the range of the target at time = 0. If the target forward velocity is zero then the start range corresponds to the constant fixed range of the target.

INPUT : Target Height
Units : metres

For simplicity the ORACLE model assumes the target can be represented by a rectangular block the height of which can be entered by selecting this item.

INPUT : Target Width
Units : metres

For simplicity the ORACLE model assumes the target can be represented by a rectangular block the width of which is specified by this variable.

The value is 6.75 metres represents the length of a typical tank without the gun barrel.

INPUT : Target/Sensor forward velocity
Units : metres / second

This variable specifies the component of the target velocity directly towards (or away from) the observer. This should account for both observer and target velocity components.

INPUT : Target crossing velocity
Units : metres / second

This variable specifies the component of the target velocity orthogonal to the observer.

INPUT :Target Intrinsic Luminance Contrast
Units : dimensionless

The intrinsic contrast of a target object against the background is a direct measure of luminance contrast.

INPUT :Background Luminance
Units : cd/m^2

This variable corresponds to the ambient luminance of the scene and is used to set the level of adaptation of the eye.

INPUT : Visibility in km
Units : kilometres

The visibility is the meteorological parameter representing the atmospheric attenuation of contrast down to the 2 % level.

INPUT : Sky to Ground Luminance Ratio
Units : dimensionless

The sky to ground luminance ratio is used in the calculation of the scattering term of contrast attenuation with Range. The higher the sky luminance relative to the ground the greater is the veiling light level and therefore contrast reduction. N.B. where the target is assumed to be viewed against a sky background this parameter must be set to 1.

INPUT : Sight Veiling Glare *
Units : dimensionless

The sight veiling glare is a measure of full field added light as a fraction of background luminance. A typical optical sight has a veiling glare in the range of 10-20 % which would be entered into the model as 0.10 or 0.20 .

INPUT : Sight Transmission (0-1) *
Units : dimensionless

This is the transmission of light as a fraction of the input energy within the photopic spectral waveband. A typical value for a multi-element sight may be as low as 12% which would be entered as 0.12.

INPUT :Diameter of Circular Field of View *
Units : degrees

The eye-space field of view is used to define the area of search. The value is required in degrees.

INPUT :Total System Magnification *
Units : dimensionless

This variable represents the total magnification of the optical system and is used in calculating the size of the image at the cornea.

INPUT : Slew Rate of Sensor
Units : degrees/second

The average rate of coverage of an arc of responsibility in degrees per second is used to model slewing performance.

INPUT : Area for Slewing Search
Units : degrees squared

The arc over which the observer slews the sight is represented as an area in square degrees in object space. The slewing model covers this area progressively until it is completely covered.

INPUT : Number of Samples in Optical MTF *
Units : dimensionless

This defines the number of samples in the Optical MTF array.

INPUT : Frequency Increment of Optical MTF *
Units : cycles per mrad

This defines the frequency increment of the entries contained within the optical sight MTF array. This variable should be set as required before any attempt to alter the entries in the MTF array.

INPUT : Optics MTF Array *
Units : dimensionless

This array describes the MTF of the optical system. No constraints are imposed on the values which may be entered into this array and the user must ensure that values are within the valid range of 0.0 to 1.0.

INPUT : Frequency of Sinusoidal Vibration *
Units : Cycles per second (Hz)

The model can account for the effects of a single vibration frequency at the eye. It does not allow for damping that occurs between the seat and eye (high frequency vibration is significantly attenuated by the neck and spine).

INPUT : Sensor Vibration Amplitude *
Units : mrads (at the cornea)

This represents the peak to peak amplitude of the vibration in mrads at the cornea.

INPUT : Target/Background Temp. difference *
Units : degrees Celsius

This represents the averaged difference in temperature between a target and the background against which it is viewed. For simplicity all temperatures are assumed to be apparent i.e. all objects have unity emissivity.

INPUT : Elevation Field of View at the Eye *
Units : degrees

This is the eye-space field of view height for a thermal imager. This assumes the shape is a rectangle and the width is the horizontal angle to the observers eye in degrees.

INPUT : Azimuth Field of View at the Eye *
Units : degrees

This is the eye-space field of view width for a thermal imager. This assumes the shape is a rectangle and the width is the horizontal angle to the observer's eye in degrees.

INPUT : Telescope Magnification *
Units : dimensionless

This allows specification of the magnification of the telescope on the thermal imager. The use of this variable assumes that the imager comprises a scanner (with fixed optics) plus an optional telescope of the desired magnification. If the system under consideration is not of this type (ie scanner and optics are integrated) then the simplest user option is to set this parameter to unity. In these circumstances variables in Variable Menu Part 2 referring to the scanner will now require values pertaining to the objective optics.

INPUT : Temperature Window (Gain) *
Units : degrees Celsius

The required temperature window (gain) in degrees Celsius is the difference in object space temperatures corresponding to black level and peak white video voltages which includes the effect of the telescope transmission.

INPUT : Air Temperature
Units : Kelvin

The ambient air temperature near ground level.

INPUT : Ground Temperature
Units : Kelvin

The average apparent ground temperature in Kelvin.

INPUT : Lower Band Limit *
Units : microns

The lowest wavelength of the detector's response.

INPUT : Upper Band Limit *
Units : microns

The upper wavelength of the detectors response.

INPUT :Wavelength Increment *
Units : microns

This defines the steps in wavelength between the lower and upper ends of the spectral band used for characterising the system.

INPUT : Scanner Field of View in Azimuth *
Units : degrees

This is the azimuth field of view of the active scan in degrees. Note that this is the field of view without the telescope.

INPUT : Scanner Focal Length *
Units : metres

The effective focal length of the scanner optics (ie excluding any telescope). It is assumed that the telescope is designed to maintain scanner f number.

INPUT : Scanner Aperture *
Units : metres

The aperture of the scanner optics alone.

INPUT : Telescope Aberration Factor *
Units : dimensionless

This factor is a power applied to the optics MTF to account for
aberration. A value of unity gives diffraction limited performance.

INPUT : Optics Transmission Array *
Units : dimensionless

The combined transmission of the scanner and telescope optics. The number of values required is calculated from the selected spectral band and wavelength increment. The number of samples in the array and the frequency interval are defined by the variables 'Lower Band Limit', 'Upper Band Limit' and 'Wavelength Increment'.The values of all these variables should be set before any attempt to edit the contents of the array.

INPUT : Discrete or Sprite *
Units : dimensionless

Option to choose between implementing Sprite or discrete detectors.

INPUT : Detector Size in X *
Units : microns

The detector width in microns. This variable is only of significance for a discrete detector.

INPUT : Detector Size in Y *
Units : microns

The detector height in microns. This variable is only of significance for a discrete detector.

INPUT : Detector Readout Length *
Units : microns

Strictly speaking the size of that section of the detector from which the accumulated charge is 'read-out'. Here the term is used in the wider sense of that dimension which gives an appropriate 'sinc' term for the two component MTF of the detector under consideration. This variable is only of significance for a sprite detector.

INPUT : Detector Diffusion Length *
Units : microns

Strictly speaking the distance travelled by charge carriers during their lifetime in the detector under given operating conditions. More broadly here we imply the dimension required to give a suitable 'diffusion' term for the detector MTF.

INPUT : "Noise Readout Length" *
Units : microns

The Sprite noise power spectrum is of similar form to the detector MTF, essentially comprising two terms. This variable is that dimension which gives an adequate fit to the 'sinc' term and by analogy with the MTF but to distinguish from it is here called the 'noise read-out length'.

INPUT : "Noise Diffusion Length" *
Units : microns

That dimension which determines the diffusion component of the Sprite noise power spectrum.

INPUT : Peak Wavelength *
Units : microns

This is the wavelength at which the peak of the detector response occurs.

INPUT : Number of Detectors in Series *
Units : dimensionless

The number of detectors in series.

INPUT : Scan Velocity *
Units : metres / second

The image velocity at the detector.

INPUT : Specific Detectivity *

Units : $m \sqrt{\ } Hz\ W^{-1}$

The peak specific detectivity at the detector f/number and temperature. If quoted in another form e.g. D*500 it will need conversion before it can be meaningfully used in the model.

INPUT : Relative DStar Array *
Units : dimensionless

This is the relative response of the detector across the selected spectral band in the selected increments. The number of samples in the array and the frequency interval are defined by the variables 'Lower Band Limit', 'Upper Band Limit' and 'Wavelength Increment':

INPUT : Freq.Interval in Scanner Space *
Units : cy/mrad

This specifies the frequency increment for the MTF (cycles per mrad) in scanner space. This increment must be used in specifying all MTF information.

INPUT : Number of Samples in Thermal MTF *
Units : dimensionless

INPUT : Boost *
Units : dimensionless

Option to include electronic boost into the MTF.

INPUT : Electronics MTF Array *
Units : dimensionless

This is the MTF of the electronics of the thermal imager at the selected scanner space frequencies.

INPUT : Boost MTF Array *
Units : dimensionless

Only available when BOOST is selected in the menu, this is the MTF of the optional high frequency emphasis or boost provided by some thermal imager designs..

INPUT : Peak Display Luminance *
Units : candelas per square metre

The peak luminance available from the display under current control settings. This may not correspond to the peak luminance available from the CRT.

INPUT : Resting Level Luminance *
Units : candelas per square metre

The resting level of Black level luminance of the display corresponding to minimum video signal input.

INPUT : Power of Luminance to Voltage *
Units : candelas per square metre per volt

The slope of the log voltage - log luminance curve.

INPUT : Display 50 % MTF frequency *
Units : cycles per picture width

This is the display 50% MTF frequency in cycles per picture width.

INPUT : Display Frame Rate *
Units : Hertz

Frequency of refresh of the display in cycles per second.

2.1.2 Raw Data

A typical set of inputs for the test case is given below:

Fixation or Glimpse Time........................ 0.333 sec
Maximum Number of Glimpses for Search........... 50.
Viewing..binocular
Confidence Level................................ 2.8
Fractional Perimeter............................ 1.
Start Range.................................... 3000. m
Target Height.................................. 1.98 m
Target Width................................... 6.75 m
Target/Sensor forward velocity.................. 15. m/s
Target crossing velocity.................. 0. m/s
Target/Background Temp. difference.............. 3.0° K
Elevation Field of View at the Eye............. 18. °
Azimuth Field of View at the Eye.............. 24. °
Telescope Magnification......................... 4.
Slew Rate of Sensor............................ 0. °/s
Area for Slewing Search......................... 200.0 °
Atmospheric Extinction Coeff.................... 0.1 / km
Air Temperature................................ 283. ° K
Ground Temperature.............................. 283. ° K
Lower Band Limit................................ 10. microns
Upper Band Limit................................ 14 microns
Wavelength Increment............................ 0.8 microns
Scanner Field of View in Azimuth................ 60.0 °
Scanner Focal Length............................ 0.03 m

Scanner Aperture................................ 0.015 m
Telescope Aberration Factor..................... 1.5
Discrete or Sprite..............................sprite
Detector Readout Length......................... 30. microns
Detector Diffusion Length....................... 25.0 microns
"Noise Readout Length".......................... 5. microns
"Noise Diffusion Length"........................ 10. microns
Peak Wavelength................................. 9. microns
Number of Detectors in Series................... 1.
Scan Velocity................................... 105 m/s
Specific Detectivity............................ 2.0E+0009 $m\sqrt{Hz}/W$
Freq.Interval in Scanner Space.................. 0.05 cy/mrad
Number of Samples in Thermal MTF................ 25.
Boost ?...no
Peak Display Luminance......................... 200. cd/m^2
Resting Level Luminance........................ 10. cd/m^2
Power of Luminance to Voltage.................. 3. cd/m^2
Display 50 % MTF frequency...................... 240. c/pic.width
Display Frame Rate.............................. 30. Hz
Temperature Window (Gain)....................... 5. °C

In the walkthough, the temperature window is varied from 5 to 40 degrees in increments of 5 degrees.

2.2 Walkthrough

The following steps describe how to compare visual aquisition performance with the imager set to different gains for the same target acquisition task, using an iterative function of the model.

1) Load the model by typing "oracle" in the relevant directory.

2) Press Return when the startup banner is shown, and "y" to return the variables to their most recent setting.

3) Select the option Thermal Imager Model from the startup menu.(Figure1)

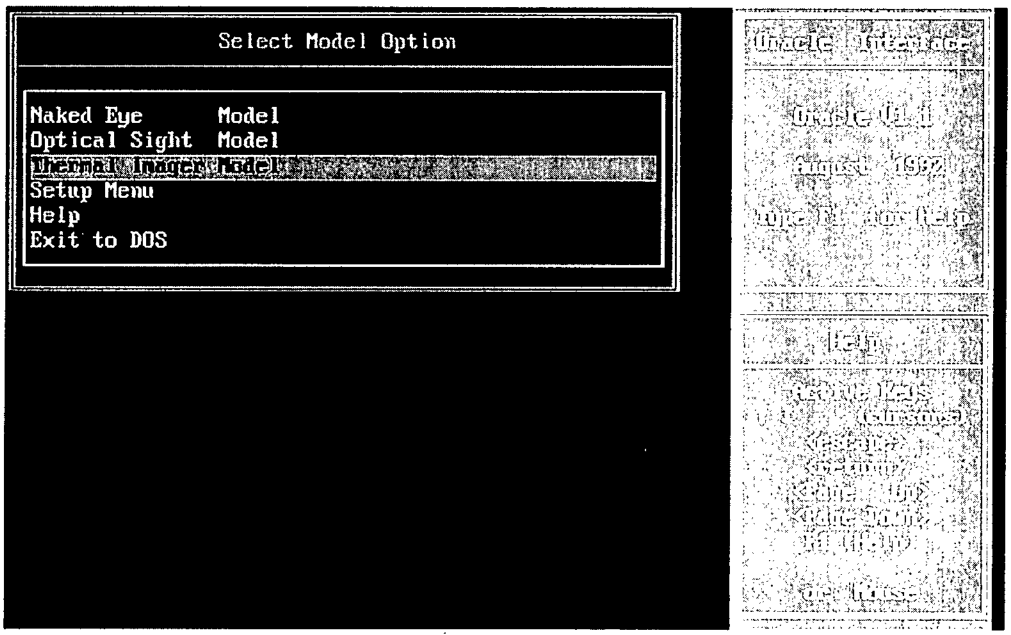


Figure 1.

4) Select the Utility menu, and in the file manipulation section, choose :

"model output data written to: both (toggled via "enter" key)

"output details : full (toggled via "enter" key")
as shown in Figure 2, then press "escape" to return to TI menu screen.

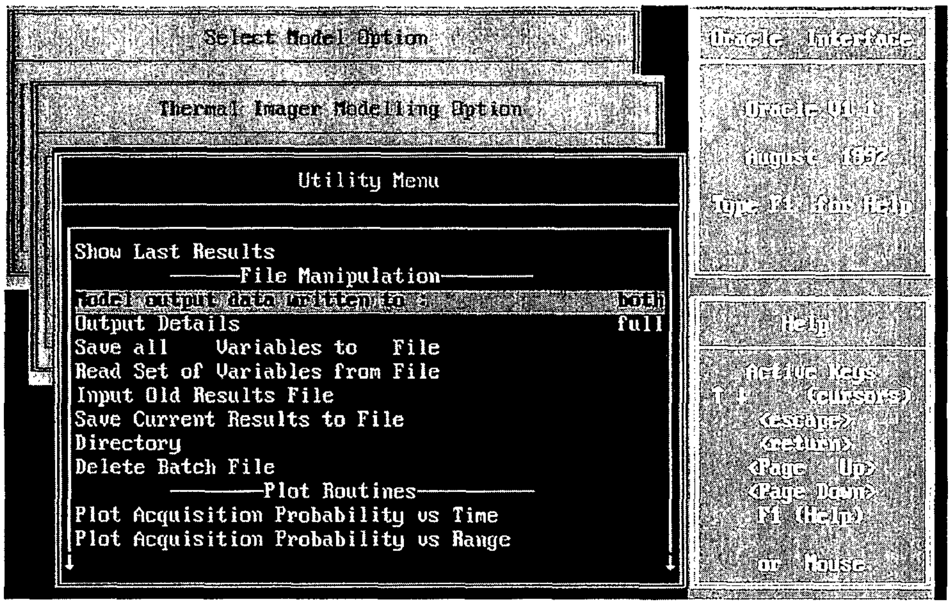


Figure 2.

5) Select the option Iterative Run of Model.(Figure 3). Scroll down to variable 'Temperature Window (Gain)' (Figure 4) and press "enter".

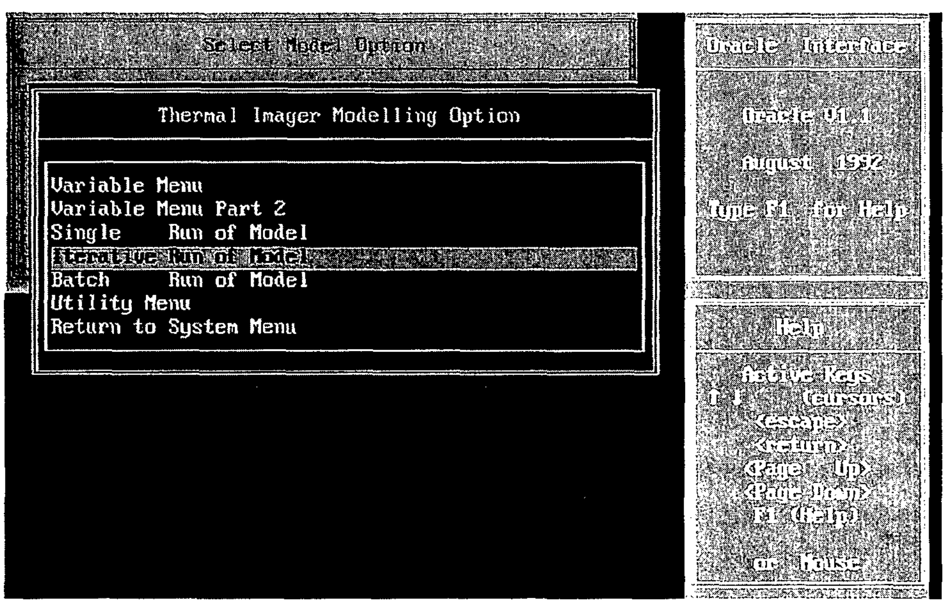


Figure 3

6) When asked to specify values input:

"iterate upwards from" 5 (enter)

"iterate upwards to" 40 (enter)
add/iteration" 5 (enter)

leave the last option blank and press "enter" (Figure 5)

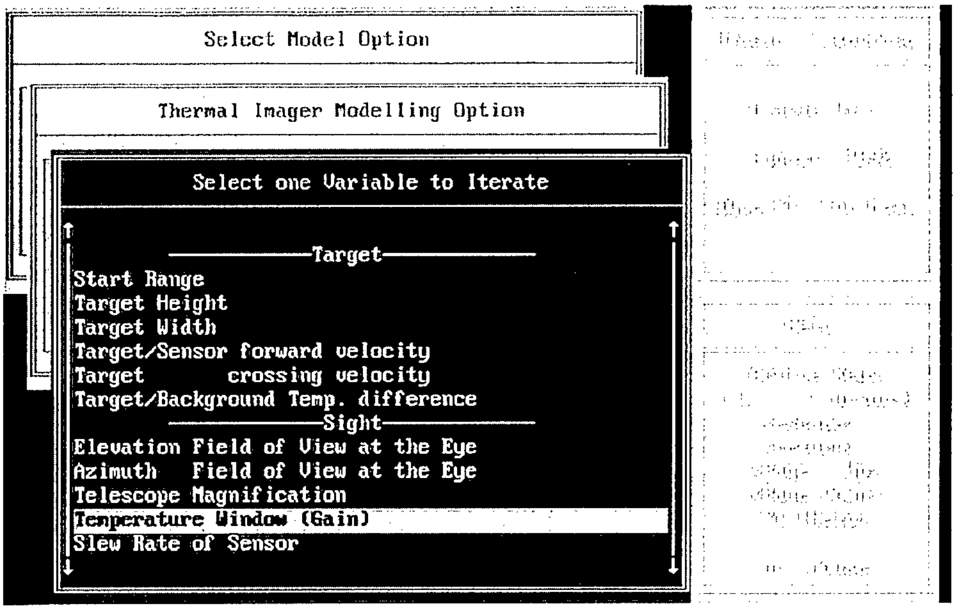


Figure 4

The model will ask for a filename - enter any valid text string.

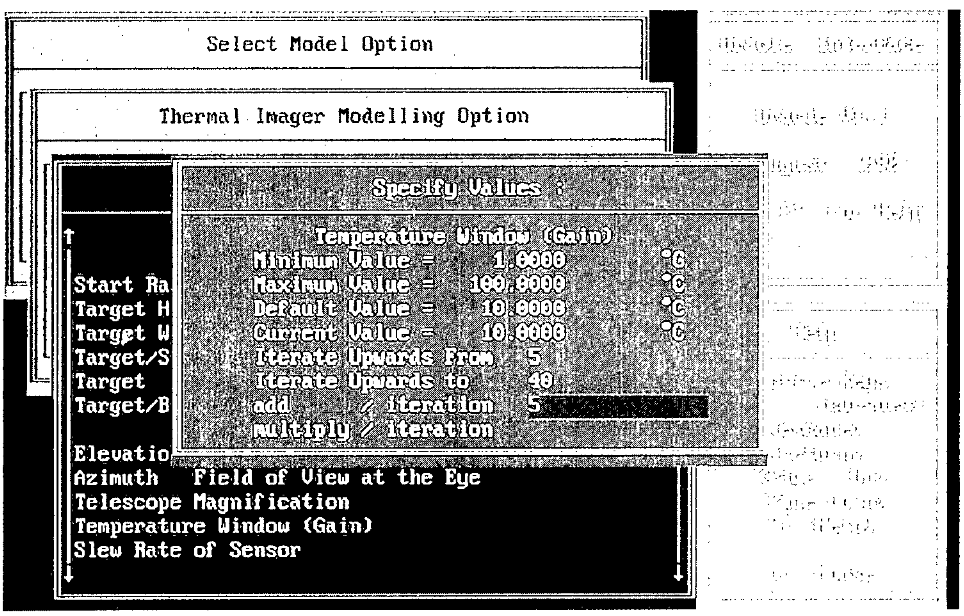


Figure 5

7) The model will run through the iterations and produce a graphical output of the aquisition probablity associated with each gain option. Press "enter" to continue.You will be asked if you wish to see the output data. As they are saved to file this is not essential. At this point the default graph is for target acquisition probability against time.

8). If you wish to see alternative data plots, go to the Utility menu and select, for example, the "plot aquisition probablity vs range" option Figure 6)

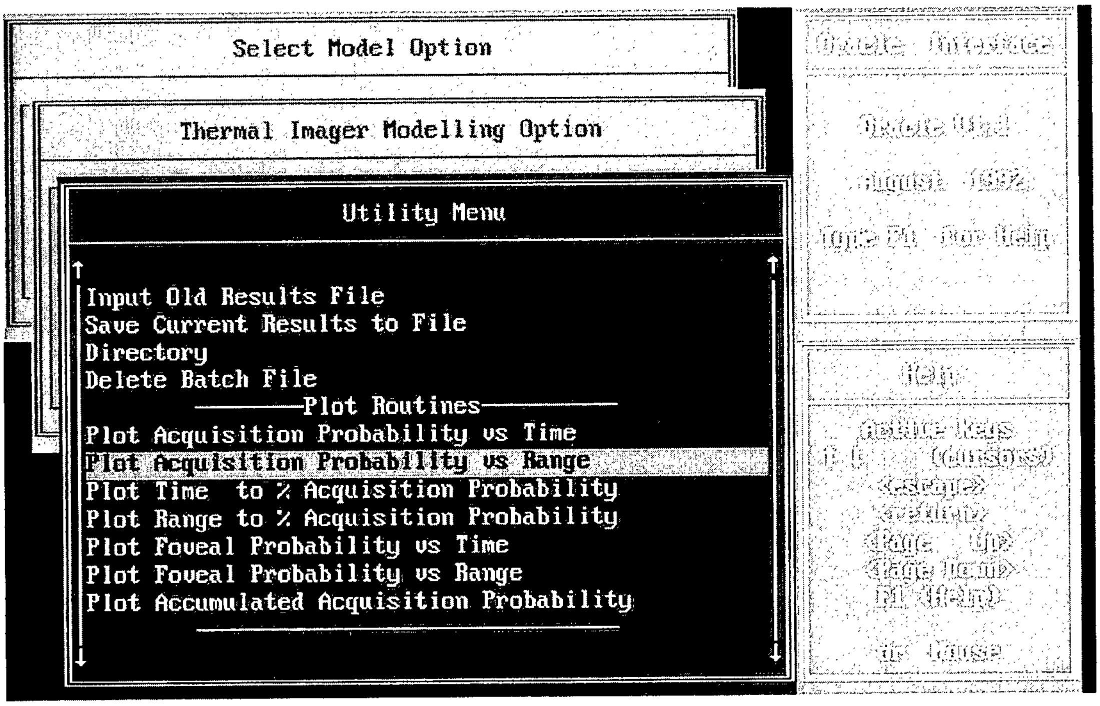


Figure 6

The saved data file is in ASCII format and can be loaded into a word processor or spreadsheet for further analysis.

2.3 Outputs

Three examples of outputs are shown below. First, there is a plot of acquisition probability versus time (Figure 7), second a plot of aquisition probability with range (distance) (Figure 8), and finally there is a partial report from the saved data file (the complete report runs to many pages). The data in the report show the performance associated with the gain set to 5 degrees and part of the performance with a gain of 10 degrees, as well as some of the underlying data for the Imager. It should be noted that the output contained in the data file contains a fuller specification of visual performance than that presented graphically - for example it includes visual performance away from the fovea.

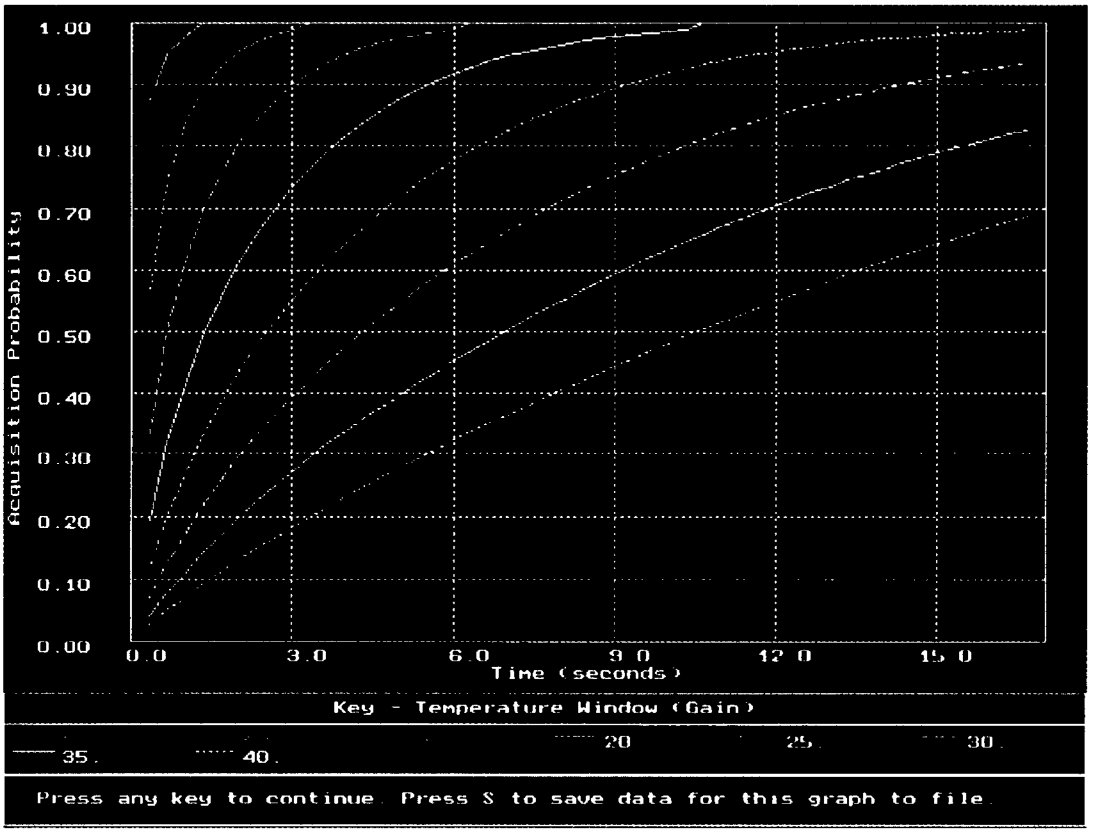


Figure 7

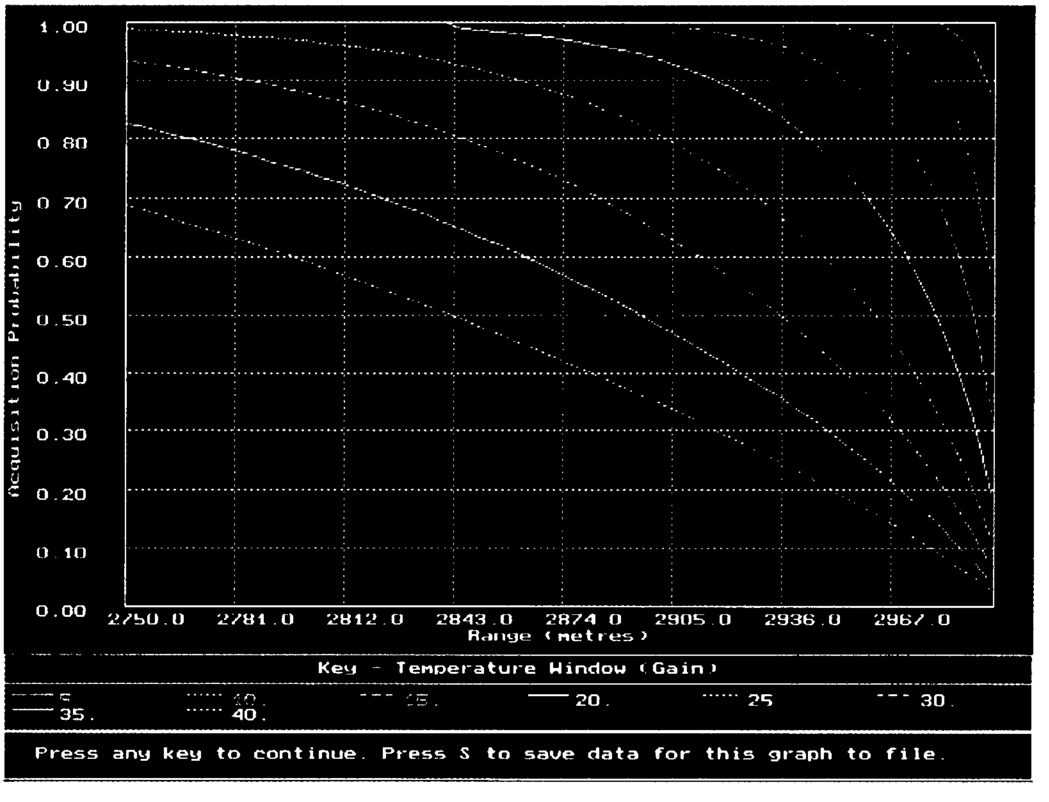


Figure 8

PARTIAL DATA OUTPUT FROM CASE STUDY

----------------------Calculated Variable Settings(partial data set only)----------------------

Point spread function width (mrads object space) 0.8467
Target radiance____________________ 7.6337 W/m²/sr/µ
Ground radiance____________________ 7.2724 W/m²/sr/µ
Sky radiance____________________ 7.2724 W/m²/sr/µ
Intrinsic radiance contrast____________________ 0.0497
Apparent radiance contrast____________________ 0.0368
Apparent target temperature____________________ 285.2314 K
Apparent grouund temperature____________________ 283. K
Background Luminance____________________ 57.5 cd/m²
Target display luminance____________________ 118.9093 cd/m²
Display contrast____________________ 1.068

------------------Visual Efficiency Across Retina------------------

Angle (°)	0	1	2	3	4	5	6
	0.432	0.372	0.334	0.308	0.287	0.271	0.257
Angle (°)	7	8	9	10	11	12	13
	0.245	0.235	0.226	0.218	0.211	0.205	0.199
Angle (°)	14	15	16	17	18	19	20
	0.193	0.189	0.184	0.180	0.176	0.172	0.169
Angle (°)	21	22	23	24	25	26	27
	0.165	0.162	0.160	0.157	0.154	0.152	0.149
Angle (°)	28	29	30	31			
	0.147	0.145	0.143	0.141			

Noise bandwidth____________________1362748.26 Hz
NETD____________________ 0.2708 °C
TI noise integration area ____________________ 0.712 mrad²(eye spce)

-------------------------Lobe Probabilities-------------------------

Angle (°)	0	1	2	3	4	5	6
	0.999	0.998	0.996	0.992	0.988	0.982	0.975
Angle (°)	7	8	9	10	11	12	13
	0.966	0.954	0.939	0.920	0.896	0.864	0.833
Angle (°)	14	15	16	17	18	19	20
	0.810	0.786	0.761	0.736	0.709	0.683	0.656
Angle (°)	21	22	23	24	25	26	27
	0.628	0.601	0.574	0.547	0.521	0.495	0.470
Angle (°)	28	29	30	31			
	0.446	0.422	0.399	0.377			

Accumulated probability = 0.8765 Range = 2995.0050 metres

Target radiance____________________ 7.6337 W/m²/sr/µ
Ground radiance____________________ 7.2724 W/m²/sr/µ
Sky radiance____________________ 7.2724 W/m²/sr/µ
Intrinsic radiance contrast____________________ 0.0497
Apparent radiance contrast____________________ 0.0368
Apparent target temperature____________________ 285.2325 K
Apparent grouund temperature____________________ 283. K
Background Luminance____________________ 57.5 cd/m²
Target display luminance____________________ 119.0379 cd/m²
Display contrast____________________ 1.0702

-------------------------Lobe Probabilities-------------------------

Angle (°)	0	1	2	3	4	5	6
	0.999	0.998	0.996	0.993	0.988	0.983	0.975
Angle (°)	7	8	9	10	11	12	13
	0.966	0.955	0.940	0.922	0.898	0.866	0.836
Angle (°)	14	15	16	17	18	19	20

0.813 0.789 0.765 0.739 0.713 0.687 0.660
Angle (°) 21 22 23 24 25 26 27
0.633 0.606 0.579 0.552 0.526 0.500 0.475
Angle (°) 28 29 30 31
0.450 0.427 0.404 0.382

Accumulated probability = 0.9489 Range = 2990.0100 metres

Target radiance	7.6337 $W/m^2/sr/\mu$
Ground radiance	7.2724 $W/m^2/sr/\mu$
Sky radiance	7.2724 $W/m^2/sr/\mu$
Intrinsic radiance contrast	0.0497
Apparent radiance contrast	0.0369
Apparent target temperature	285.2336 K
Apparent ground temperature	283. K
Background Luminance	57.5 cd/m^2
Target display luminance	119.1668 cd/m^2
Display contrast	1.0725

-------------------------Lobe Probabilities-------------------------
Angle (°) 0 1 2 3 4 5 6
0.999 0.998 0.996 0.993 0.988 0.983 0.976
Angle (°) 7 8 9 10 11 12 13
0.967 0.955 0.941 0.923 0.899 0.868 0.838
Angle (°) 14 15 16 17 18 19 20
0.816 0.792 0.768 0.743 0.717 0.691 0.664
Angle (°) 21 22 23 24 25 26 27
0.637 0.610 0.583 0.557 0.531 0.505 0.480
Angle (°) 28 29 30 31
0.455 0.431 0.408 0.386

Accumulated probability = 0.9783 Range = 2985.0150 metres

Target radiance	7.6337 $W/m^2/sr/\mu$
Ground radiance	7.2724 $W/m^2/sr/\mu$
Sky radiance	7.2724 $W/m^2/sr/\mu$
Intrinsic radiance contrast	0.0497
Apparent radiance contrast	0.0369
Apparent target temperature	285.2347 K
Apparent grouund temperature	283. K
Background Luminance	57.5 cd/m^2
Target display luminance	119.2961 cd/m^2
Display contrast	1.0747

-------------------------Lobe Probabilities-------------------------
Angle (°) 0 1 2 3 4 5 6
0.999 0.998 0.996 0.993 0.989 0.983 0.976
Angle (°) 7 8 9 10 11 12 13
0.967 0.956 0.942 0.924 0.901 0.870 0.841
Angle (°) 14 15 16 17 18 19 20
0.819 0.795 0.771 0.747 0.721 0.695 0.668
Angle (°) 21 22 23 24 25 26 27
0.642 0.615 0.588 0.562 0.535 0.510 0.484
Angle (°) 28 29 30 31
0.460 0.436 0.413 0.391

Accumulated probability = 0.9911 Range = 2980.0200 metres

Time (sec)	Range (m)	Static Prob.	Foveal Prob.
0.3330	2995.005	0.8765	0.99908
0.6660	2990.01	0.9489	0.99909
0.9990	2985.015	0.9783	0.99911
1.3320	2980.02	1.	0.99912

Temperature Window (Gain)...................... 10. °C

--------------------Calculated Variable Settings--------------------

Point spread function width (mrads object space) 0.8467
Target radiance_________________ 7.6337 $W/m^2/sr/\mu$
Ground radiance_________________ 7.2724 $W/m^2/sr/\mu$
Sky radiance_________________ 7.2724 $W/m^2/sr/\mu$
Intrinsic radiance contrast_________________ 0.0497
Apparent radiance contrast_________________ 0.0368
Apparent target temperature_________________ 285.2314 K
Apparent grouund temperature_________________ 283. K
Background Luminance_________________ 57.5 cd/m^2
Target display luminance_________________ 85.0648 cd/m^2
Display contrast_________________ 0.4794

-------------------Visual Efficiency Across Retina-------------------

Angle (°)	0	1	2	3	4	5	6
	0.432	0.372	0.334	0.308	0.287	0.271	0.257
Angle (°)	7	8	9	10	11	12	13
	0.245	0.235	0.226	0.218	0.211	0.205	0.199
Angle (°)	14	15	16	17	18	19	20
	0.193	0.189	0.184	0.180	0.176	0.172	0.169
Angle (°)	21	22	23	24	25	26	27
	0.165	0.162	0.160	0.157	0.154	0.152	0.149
Angle (°)	28	29	30	31			
	0.147	0.145	0.143	0.141			

Noise bandwidth_________________1362748.26 Hz
NETD_________________ 0.2708 °C
TI noise integration area _________________ 0.712 $mrad^2$(eye spce)

-------------------------Lobe Probabilities-------------------------

Angle (°)	0	1	2	3	4	5	6
	0.997	0.992	0.983	0.969	0.948	0.918	0.880
Angle (°)	7	8	9	10	11	12	13
	0.831	0.772	0.701	0.622	0.534	0.441	0.370
Angle (°)	14	15	16	17	18	19	20
	0.324	0.283	0.248	0.217	0.189	0.166	0.146
Angle (°)	21	22	23	24	25	26	27
	0.128	0.113	0.099	0.088	0.078	0.069	0.062
Angle (°)	28	29	30	31			
	0.055	0.049	0.044	0.040			

Accumulated probability = 0.5702 Range = 2995.0050 metres
Target radiance_________________ 7.6337 $W/m^2/sr/\mu$
Ground radiance_________________ 7.2724 $W/m^2/sr/\mu$
Sky radiance_________________ 7.2724 $W/m^2/sr/\mu$
Intrinsic radiance contrast_________________ 0.0497
Apparent radiance contrast_________________ 0.0368
Apparent target temperature_________________ 285.2325 K
Apparent grouund temperature_________________ 283. K
Background Luminance_________________ 57.5 cd/m^2
Target display luminance_________________ 85.1181 cd/m^2
Display contrast_________________ 0.4803

------------------------Lobe Probabilities-------------------------

Angle (°)	0	1	2	3	4	5	6
	0.997	0.992	0.983	0.969	0.948	0.920	0.882
Angle (°)	7	8	9	10	11	12	13
	0.834	0.775	0.705	0.626	0.539	0.446	0.374
Angle (°)	14	15	16	17	18	19	20
	0.328	0.287	0.251	0.220	0.193	0.169	0.148
Angle (°)	21	22	23	24	25	26	27

	0.130	0.115	0.101	0.089	0.079	0.070	0.063
Angle (°)	28	29	30	31			
	0.056	0.050	0.045	0.040			

3. SOLUTION DESCRIPTION

A solution to the design question is suggested by Figure 7. For this viewing condition, the best acquisition performance is obtained with the low gain system settings. As with most systems, however, there is a trade-off in performance between parameters. For the gain of a TI, there is a trade-off with the dynamic range available - an increase in gain often leads to increased visual noise in the display. Consequently an optimised gain for the display may also yield worse performance on other tasks (for example in scenes with large variations in brightness). Further model runs would be required to investigate these trade-offs, but the effort required is minimal.

4. FACILITY/RESOURCE REQUIREMENTS.

This worked example was performed on a version of ORACLE running on a PC under DOS. There are no special requirements of the PC, although the faster the CPU the quicker the iterative calculations can be made. Some knowledge of TI's are required if changes are to be made to the default values, and a working knowledge of basic photometric terms is helpful in understanding the visual parameters. The model run took approximately 1 hour to set up and document, with help from the Vision Group at BAe. Most of this time was devoted to documentation - preparing the inputs to the model and the running time take about 25 minutes.

Case Studies Involving W/Index

Dr. Christopher A. Miller
Honeywell Technology Center—MN65-2600
3660 Technology Dr.
Minneapolis, MN. 55418
U.S.A.

cmiller@htc.honeywell.com

1. Summary

This document describes in detail the capabilities of Honeywell's Workload Index (W/Index) tool, its assumptions and philosophy, methods of use, and types and utility of output. Two case studies are provided to illustrate the process and applicability of workload prediction using W/Index: (1) an example evaluating crew station layout and functionality in an advanced attack/scout helicopter domain, and (2) an example evaluating alternate methods of crew reduction through added automation in an existing tank.

2. Problem Space

We report on the use of a human resource-based simulation tool, the Workload Index (W/Index) to initiate performance evaluations of alternate crew station designs very early in the design cycle. This tool uses a multiple resource model [1] of human attention to represent the levels of conflict a human operator incurs when performing tasks in a hypothetical crew station. While similar to workload-based crew station evaluation, our approach differs in that it is grounded in the physical layout of the proposed cockpit and the physical capacities of a human operator, rather than in abstract or subjective notions of workload. Also, we use our methodology for initial design guidance rather than for later evaluation (e.g. TLX, SWAT). Results show an extremely rapid capability to study performance effects of alternate crew station design.

The design of crew stations is often a process of generate and test. Designers generate conceptual crew stations (in whole or part) which are then reviewed and tested by end users (e.g., pilots) to assess their acceptability, safety, and effectiveness. Pertinent data can generally only be collected via human-in-the-loop interaction with a crew station prototype, and higher fidelity prototypes generally provide richer, more detailed and more accurate data. Unfortunately, human-in-the-loop testing, especially with high fidelity prototypes, is costly and time consuming. For these reasons, in traditional design approaches (e.g., [2]), human performance testing is a serious bottleneck.

This situation forces most crew station design efforts to be conservative. Departures from traditional designs are rare and small—both because existing designs are *known* to be acceptable and because greater deviations will require increased testing. Once testing is begun on a prototype, there can be substantial resistance to change. The reasons for this stem from the nature of the testing itself. Traditionally, the only valid measures of successful crew station design have been operator acceptance and adequate human-system performance. To obtain data for these measures, a substantially complete design has to first be composed and then implemented in a human-in-the-loop prototype. Not only does this require substantial upfront costs (thereby making redesign, and retesting, unlikely), it also makes it extremely rare for multiple, candidate designs to be developed and tested against each other. Thus, the traditional design approach may produce an *acceptable* crew station, but there is no way of knowing whether or not it is there might be a better one.

3. Description of Process

We have developed a human performance simulation tool to push aspects of human-in-the-loop performance testing much earlier in the design cycle. This tool, the Workload Index (W/Index-- [3]), enables a coarse-grained simulation of the human performance impact of many important aspects of human-machine system design—crew compliment, automation behaviors, operator task loading, operational procedures, display and control design, etc.—long before a design is complete, much less before a human-in-the-loop prototype can be constructed. In the remainder of this paper, we briefly describe W/Index and then present our method of using it in early crew station design. Finally, we provide some illustrative results from a crew station design effort in an advanced attack/scout helicopter domain, and from a crew reduction study in an existing tank.

3.1 *The W/Index Modeling Tool*

The Workload Index (W/Index) tool was developed by Honeywell to predict operator workload due to the conflicts incurred by multiple concurrent tasks making simultaneous use of the same human

resource. W/Index is designed to provide relative measures of the conflict levels produced by alternate crew station designs over the course of one or more representative mission scenarios. W/Index allows system designers to consider the taskload consequences of the physical layout of the crew station, the application of automation to crew tasks, the use of various human-machine interface technologies, and the sequence of crew task loading. W/Index implements a workload estimation algorithm based on Wickens' Multiple Resource Theory [1] modeling resource demands on a single human operator performing a static task timeline. W/Index has been used to perform taskload and conflict analysis for projects in both military and commercial aviation, as well as crew reduction studies for U.S. Army tanks.

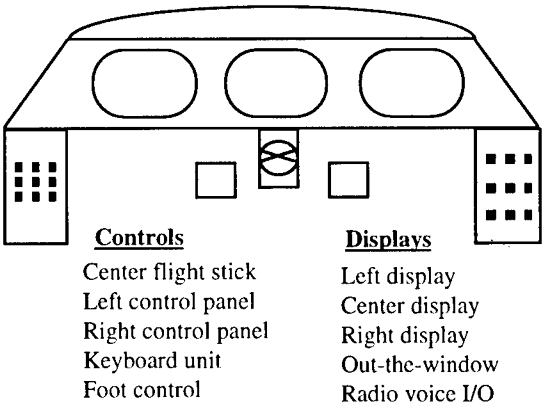


Figure 1. A Conceptual Crew Station.

3.2 *Conceptual Design Process*

To use W/Index to evaluate any crew station, automation, procedure or interface design, five components are needed:

1. Multiple concepts to be evaluated against each other (e.g., alternative interface designs, crew station layouts or automation behaviors).
2. Mission scenarios with tasks (and their sequential relations) to be performed by human operator(s) with the crew station.
3. An Interface Activity Matrix which defines which crew resources will be used for the performance of each task in the timelines.
4. A Conflict Matrix defining the degree of resource conflict whenever two or more attentional resources are required simultaneously to perform one or more tasks.
5. An algorithm for calculating conflict levels throughout the timelines (provided in W/Index itself).

Each of these components will be described in more detail below.

3.2.1 Candidate Crew Station Concepts

Conceptual crew stations for evaluation via this methodology need only be developed as lists of controls and display channels, resource usages and attentional demand levels. W/Index can provide data on the resource demands of a design at various levels of "granularity." If the design is in its early phases, it is not necessary to consider the formats of the information displays, the exact location of screen bezels or stick buttons, etc. However, if the design is near completion and the location and behavior of controls and displays are well defined, a higher level of detail can be used. In either case, since W/Index provides data about the resource demands of one design relative to another, it is important that both designs be modeled at the same level of granularity.

A hypothetical cockpit with adequate detail for our analyses is illustrated in Figure 1. We recently used W/Index in the very early design phases of a dual-crew, advanced attack/scout helicopter whose crew station had been designed to approximately this level of detail (i.e., a general physical layout of controls and displays but no specific articulation of display formats or operations).

3.2.2 Critical Mission Segments

Once the crew station's physical layout is determined, a series of "critical mission segments" are developed to simulate interaction of the operators and crew station in high workload, high criticality conditions. Most of the crew's mission time consists of redundant, comparatively low workload/low criticality tasks, but these short (2-3 minute) segments are chosen to represent "worst case scenarios" for crew operations. In general, it is unnecessary to model a full mission; optimizing the crew station for these critical mission segments will improve overall mission success and human-machine performance.

For the advanced attack/scout helicopter, one critical mission segment simulated was a battle handover. Here, operators must not only safely maneuver the helicopter and detect and carry out actions with regards to an enemy, but also coordinate their maneuvers and communications with incoming friendly helicopters, all in a rapidly changing, high-threat environment. Such an interval is critical to mission success, yet high levels of resource conflict may result from excessive verbal and visual communications, nap-of-the-earth flying, incoming auditory and visual data, etc. A well-designed crew station can minimize operator resource conflicts produced during such an interval, thereby improving operator performance; a poorly-designed crew station can increase conflict, making successful performance virtually impossible. By using CREWCUT and W/Index as modeling tools we can study predicted conflict levels and thereby, human performance effects, in a variety of candidate crew stations during critical mission segments like this one, long before commitments are made to crew station construction.

A task timeline is composed of multiple tasks or "activities", each with its interface channel requirements, which may occur once, repeatedly or continuously throughout the critical mission segment. Subject Matter Experts (SMEs) provided data for defining activities and assembling them into task timelines. For the attack/scout helicopter analysis, we first defined four critical mission segments, each containing multiple tasks for the two crewmembers and cockpit automation, as well as world events. The goal of these timelines was not strict accuracy in modeling events during mission performance, but rather to create a plausible testbench to evaluate different candidate cockpit designs. This motivation leads to many compromises in model development, as discussed below.

For many task steps, it is impossible to say when, precisely, the step will take place. This is especially true of "continuous" tasks such as those involved in flying the aircraft or monitoring aircraft subsystems (e.g., fuel status). Tasks of this nature must be done "continuously," but the physical resources used to, for example, fly the aircraft, may admit "disengagements" of up to several seconds in some circumstances (e.g., hands off stick, eyes removed from flight displays, etc.) Modeling tasks of this sort has traditionally been a problem for approaches to workload prediction, since the scheduling of these tasks is partially under operator control and permits various workload management strategies. By focusing on the problem of evaluating alternative cockpit configurations, we eliminate the need to be overly concerned with *when* these tasks are performed. Instead, we can assume an unrealistic or worst case frequency of task steps to serve as a "background" against which to evaluate conceptual crew stations. Although we know this produces an unrealistically high *absolute* estimate of conflict in the results of our simulations, as long as we use the same pattern of task steps in evaluating alternative crew stations, those designs which yield lower *relative* conflict values will generally produce better human-machine performance than those which yield higher conflict levels.

W/Index requires a static, single-path timeline (consisting only of start and stop times for all tasks or activities) for a single operator. The timeline may (in fact, it is expected to)

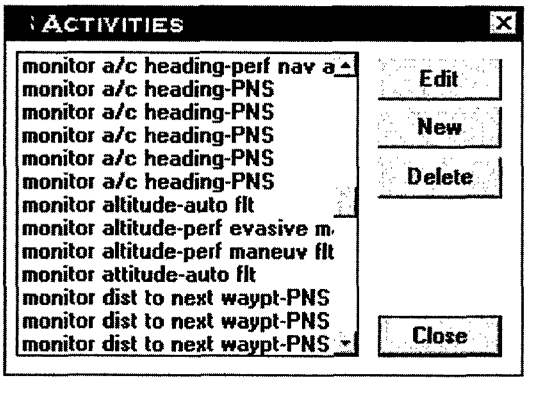


Figure 2a. W/Index list of previously defined activities for the helicopter scenarios.

represent the performance of multiple tasks in parallel, but unlike the partially-ordered graphs and sequential dependencies represented in MicroSAINT, or the alternative workload management strategies in CREWCUT which permit multiple paths through a task "network", W/Index permits no branching logic. Of course, multiple paths through a task network can each be modeled and run as separate task timelines with comparatively little effort in W/Index. The W/Index listing activities defined for the helicopter study is presented in Figure 2a while the screen for defining a new activity (reached by selecting "New" from the Activities screen in Fig 2a) is shown in Figure 2b. Note that the Edit Activity screen allows the definition of the activity in terms of the cockpit

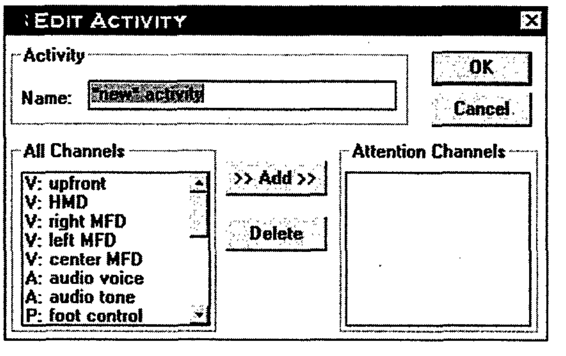


Figure 2b. W/Index Activity Definition screen.

channels which will be used whenever that activity is ongoing. The creation of channels and linking them to activities will be discussed in the next section below.

Once all needed activities have been defined, a timeline is created by assigning start and stop times for each instance of each activity which will occur during the timeline. Figure 3 shows the timeline creation and editing window in W/Index. Previously defined activities can be selected by pulling down the scrolling window in the "Edit Instance" frame, and then a start and stop time must be assigned to that instance of the activity. Figure 3 shows that the activity "monitor a/c heading-perf nav" has been selected and assigned a start time of .750 seconds into the scenario and a stop time of 1.750. Note that the timeline being constructed is presented in a scrolling frame at the bottom of the Time Line window. Instances of a previously defined activity can be added or deleted from the existing timeline and, as the timeline is built or modified, it can be saved via this window.

3.2.3 Interface/Activity Matrix

Each activity must also be assigned resource channels which the human operator will be required to use whenever that task is active. Resource channels

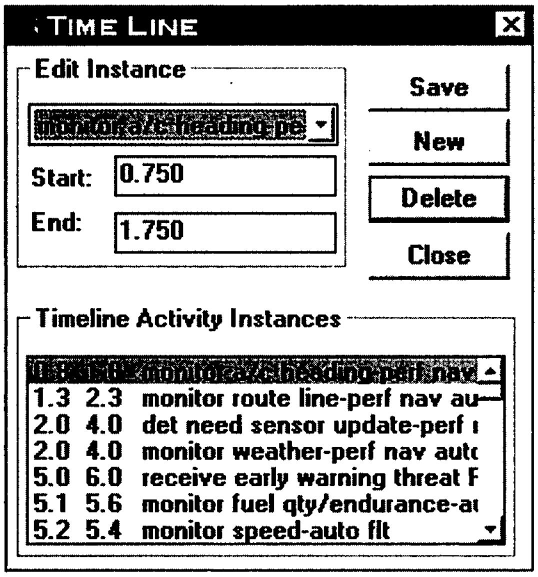


Figure 3. W/Index Time line construction window.

correspond to the physical interfaces present in the cockpit, plus human cognitive channels. Some tasks may require only a single channel (e.g., check radar status: Right display), while others may require several channels (replan route: Center display, Left display, Center control panel, keyboard unit, and spatial cognition). Alternative channels for activities can be regarded as alternative cockpit designs and may be tested against the primary channels in separate W/Index runs to evaluate predicted crew performance differences.

Figure 4a shows the list of previously defined cockpit channels for the helicopter scenarios, while Figure 4b shows the screen for creating or editing channels. When a new channel is defined, it must be assigned to one of six attentional categories: visual, auditory, kinesthetic, psychomotor, speech or cognitive. These are the only categories currently supported by W/Index, though these could be revised by interaction with the source code. Note that when a new channel is added, a "conflict value" must be assigned for the degree to which that channel conflicts with all other cockpit channels. This value will be discussed in more detail in the following section.

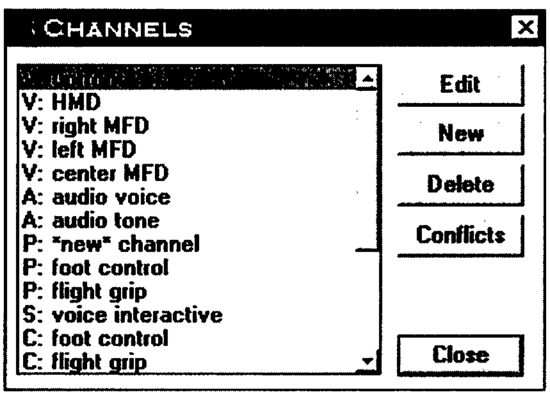


Figure 4a. List of previously defined channels in this W/Index scenario.

In previous versions of W/Index it was also necessary to indicate the degree of attention, on a five-point scale, which was demanded by each resource channel for the task. This was similar to the Aldrich [4] method of representing attentional demand levels. While this approach has conceptual appeal since it

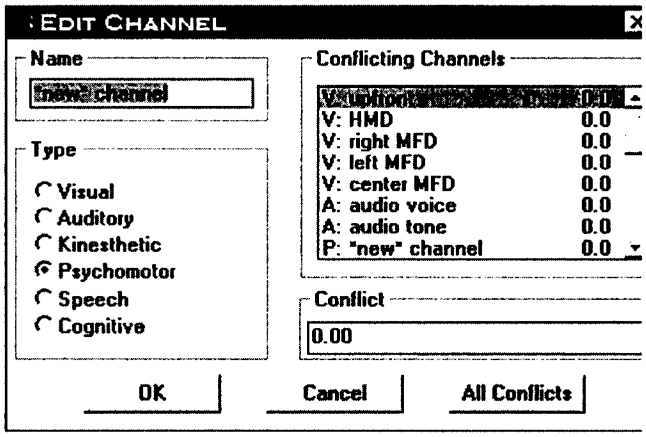


Figure 4b. Channel Definition window in W/Index.

allows us to differentiate between the *degrees* to which tasks use up the capability of a given resource channel (e.g., vision, left/right manual, etc.), it was extremely time consuming and prone to between-subjects variations. Recent work by Riley [5] has shown that attentional demand levels add no significant benefit to the predictive power of the conflict calculations for evaluating workload effects based on the placement of information. For this reason, they have been eliminated from the current version of W/Index. Recent work, however, suggests that they may still be useful for evaluating workload effects derived from automation usage and important in driving an adaptive automation scheme. Thus, we may provide them as an optional input in future W/Index versions.

3.2.4 Conflict Matrix

The final component of the modeling approach is a "Conflict Matrix" representing the degree of conflict between each pair of resources in the conceptual cockpit on a scale from 0 (essentially no conflict) to 1 ("total" conflict—these two activities cannot be done simultaneously). The values in the Conflict Matrix should be constructed using the guidelines of Multiple Resource Theory [1]. In brief, this theory claims, with support from dual-task experiments, that two simultaneous tasks which draw on the same "pool" of attentional resources will be performed less well than two tasks which draw on different resources. The set of resource "pools" consists of, roughly: vision,

audition, motor, speech, and cognition. The conflict matrix instantiates Multiple Resource Theory by, for example, ensuring that any two visual tasks receive a higher conflict value (e.g., .7-.9) than any visual + auditory task pair (.2-.4).

Given these considerations, a conflict value must be assigned for every combination of pairs of channels. This may be done for a newly defined channel via the Edit Channel window presented in figure 4b above. Alternatively, all previously defined pairwise conflict values may be reviewed and edited by selecting the "All Conflicts" button on either the Channel screen (Figure 4a) or the Edit Channel screen (Figure 4b). This results in accessing the window presented in Figure 5.

4. Solution Description

4.1 *Calculating Conflict Levels*

Given the conceptual crew station, mission segments, a task/activity matrix, and a conflict matrix, the degree of conflict for each operator can be calculated at any point in the segment as the sum of all pairwise conflict values incurred by the resources required for all concurrent tasks at that time. If attentional demand values are used, then pairwise conflict values are weighted by attentional demand values. These operations are performed automatically over the timeline provided when W/Index is asked to calculate workload values for the scenario. While this equation is simpler than that used in many workload-based assessment or prediction approaches, it provides as much predictive power as any other method while providing the most useful information regarding display and control type and location (cf. [5]).

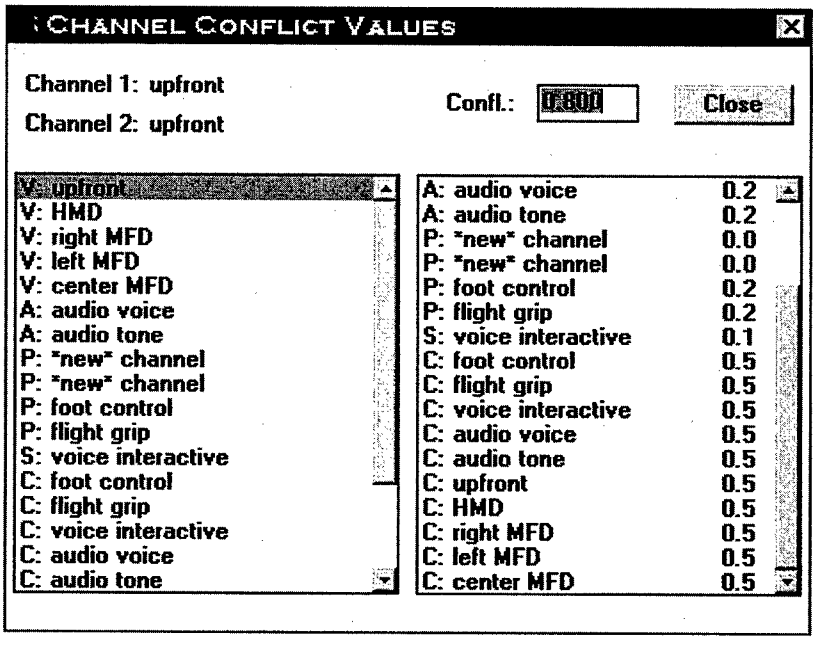


Figure 5. Pairwise Channel Conflict review and editing window.

4.2 *Using Conflict Levels in Design—An Automation Example*

W/Index provides a conflict profile for the operator throughout the timeline. The true power of W/Index, however, is in its ability to quickly assess how a change in interface or task assignments might affect the operator's workload profile. This is, therefore, a simple, low cost method of redesigning an entire crew station and ascertaining the effects on human performance. By using the baseline conflict profile for a segment, we can perform multiple permutation analyses corresponding to speculative modifications to the crew station, task allocation, or operational procedures. A conflict profile using the revised model is then compared to the baseline model and the impact of the changes analyzed.

Figure 6 presents one illustration of this approach from our scout/attack helicopter study. In this example, we envisioned a decision aid to help the pilot monitor the presence and locations of enemies and team members—that is, a piece of automation which would monitor sensor data to compare the location of team members and enemies and alert the human crewmembers of evolving threat situations. Note that this aid is far from being developed, and that one motivation for doing this permutation analysis was to decide whether such an aid would be valuable.

In the baseline model, these tasks required monitoring the Center display and using spatial cognition with reasonably high levels of attention. For the permutation modeled in Figure 6, we envisioned a smart, automated aid which would track enemy and friendly locations and movements, and alert the pilot when an unanticipated threat was evolving. Since this aid essentially enables managing these tasks by exception, we modeled no pilot resources expended for these tasks during most of the segment. Alternative (and perhaps more realistic) approaches might include an aid that provides movement projections and threat identification on a display—thereby greatly reducing the cognitive demands of these tasks while retaining most of the visual demands.

The output data from two separate W/Index runs are graphed (using Microsoft Excel's Chart Wizard) in Figure 6. These results show that the hypothetical aid produces large drops in conflict over the baseline crew station

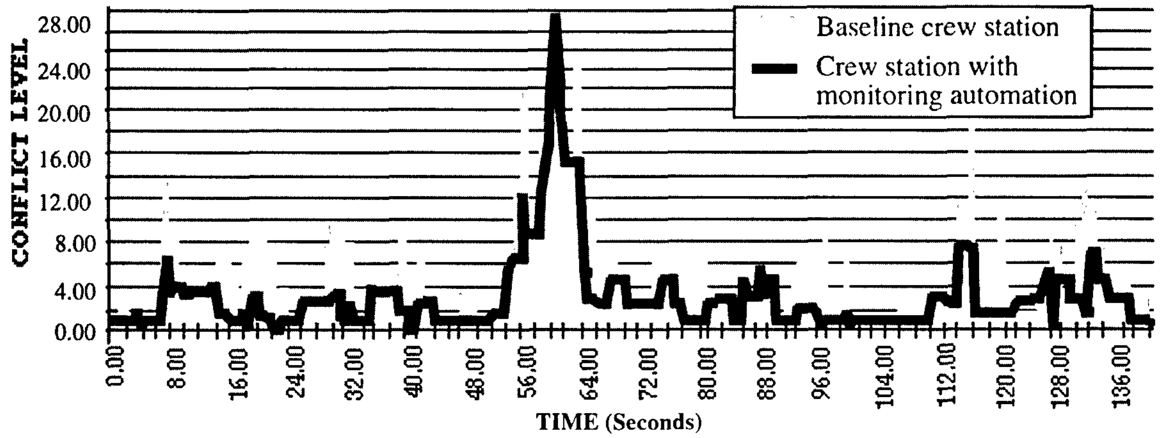
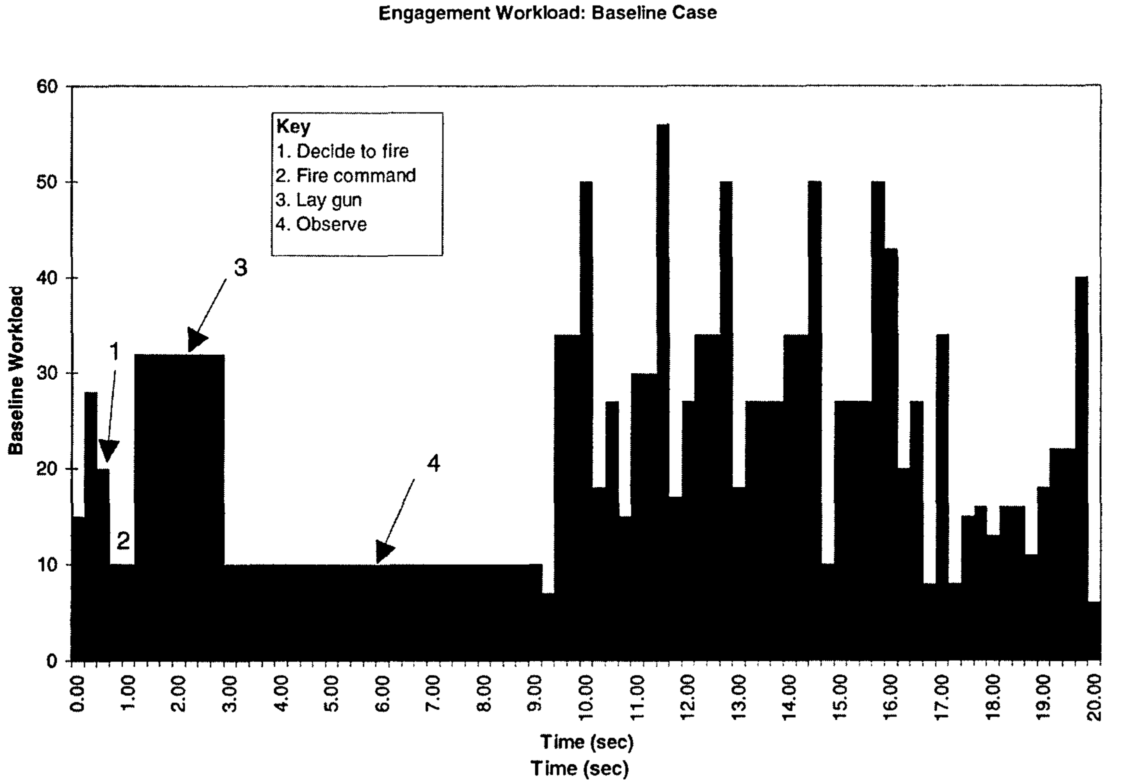


Figure 6. Sample Conflict Profiles Produced by Two Alternative Conceptual Crew Stations.

and, better yet, produces them in some of the most heavily loaded portion of the timeline. These conclusions lend weight to the belief that such an aid is a high payoff development area for the proposed cockpit. By comparing the expected payoffs of other crewstation modifications, including alternate layouts, procedures, task requirements and automation aids, we could assess relative levels of conflict reduction and provide recommendations for future resource expenditures.

4.3 *Using Conflict Levels in Design—A Crew Compliment and Task Allocation Example*

Figures 7-10 come from a program in which we applied W/Index to a crew reduction study for the Army's National Training Center (NTC). This study evaluated various automation concepts for producing a two-man version of the NTC's Opposition Force tanks (Tank Commander--TC and Driver--D but no Gunner). The mix of automation and human crew members were required to continue to perform the tasks of the former three-man crew neither significantly worse nor better than the former crews.



Time (sec)

Figure 7. TC's overall workload estimate during engagement scenario in baseline (3 crew) condition-- W/Index output.
Figure 7. TC's overall workload estimate during engagement scenario in baseline (3 crew) condition-- W/Index output.

To perform these analyses, we modeled a number of high-workload mission segments (generally engagement scenarios) totaling approximately two minutes of real-time, and then altered the task timelines and conflict matrices in W/INDEX to explore the impact of various automation concepts on performance in these segments. The example shown concerns approximately 20 seconds of the TC's overall workload estimate during an engagement scenario. We will be primarily concerned with the first 10 seconds-- in which the tank crew must identify a target, lay the gun, target the gun, fire a round, and begin to move out.

Figure 7 shows the TC's workload in the baseline, 3-man crew condition. Note that the TC has small workload peaks corresponding to deciding to fire and then laying the gun, but then is relatively unencumbered from 3 seconds until about the 10 second mark when the tank begins to move again, during which time the gunner is targeting the gun and firing it. This gap suggested that the commander could accept one or more additional tasks during this time period.

One of the gunner's tasks in tank operation is to assist the tank commander in searching for targets, effectively expanding the TC's field of view (FOV). If the gunner spots a target, he notifies the TC about it and proceeds to move to his targeting sight. If the TC spots a target, he notifies the gunner who, again, moves to his targeting sight. In either event, the TC then manually moves to gun to the approximate location of the target and issues a fire command. The gunner does precision adjustments to the gun, alerts the crew that he is about to fire and then fires the gun.

One portion of a two-man automation concept explored during this study involved the use of a sensor and display for the TC to emulate the gunner's search tasks. Sensors enabled the TC to view a 90° FOV centered around the gun via the Commander's Display (Figure 8). Additional sensors representing the gunner's FOV could be assigned by the commander to either wide (180°) or narrow (90°) modes and centered around any of the cardinal compass points (Figure 9). Since the commander's display could only present a 90° FOV, targets identified by the "automated gunner" were "pegged" to the perimeter of the display and the TC could maneuver his sensors (at the same time he was maneuvering the gun), via a joystick, to find and identify the targets. Since all sensors were slaved to the turret, moving a target into the TC's display would generally take it out of the automated gunner's FOV. Then, the TC was required to transition the automated gunner from search to engage modes (emulating the gunner's task of moving from search windows to his targeting sight) and press a fire button to enable the automated gunner to complete precision targeting and fire the gun. Following the firing, the TC was required to transition the gunner from engage mode back to search mode and reposition the gunner's sensor's to the desired configuration.

Figures 10 shows the TC's estimated workload resulting from this automation concept and crewstation design in a scenario in which a target first appears in the automated gunner's FOV. Several effects are apparent from the W/INDEX simulation. First, the task of localizing the target has become nearly 50% more difficult (in terms of relative workload scores) than it was in the baseline scenario as the TC must find the target in an unaccustomed search area. Next, laying the gun takes longer under this automation concept than it did under the baseline concept, but actually involves slightly less workload--not surprising given that the TC is interacting primarily with a visual display rather than the cedilla switch used in the baseline tank. Note also, that once the TC has laid the gun, the automated gunner can fire it almost instantaneously-- thus the tank crew can get a shot off in 6-7 seconds under this automation concept as compared to nearly 10 seconds in the baseline concept. Finally, the need to reposition the sensor at the end of the firing sequence, roughly coinciding with the need to begin moving the tank again, greatly increases the TC's workload at the end of the sequence. This is a relatively high peak and may cause workload problems in some instances.

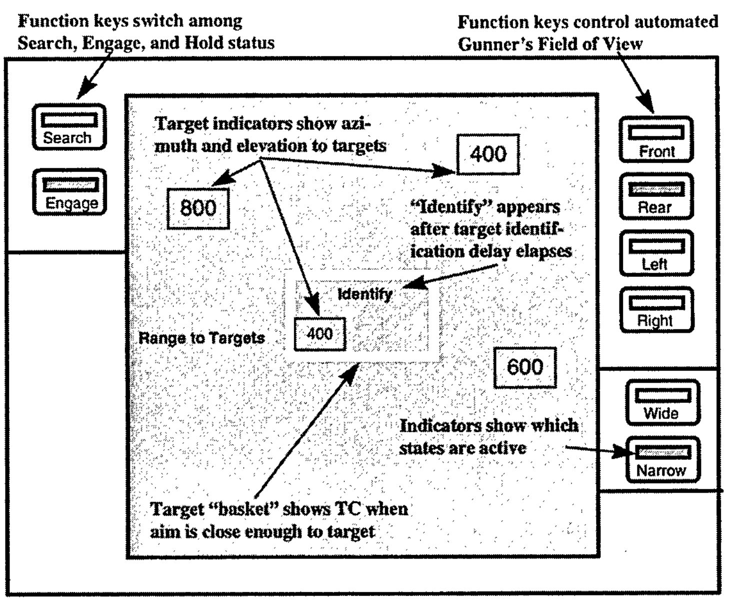


Figure 8. Proposed TC display and control interface.

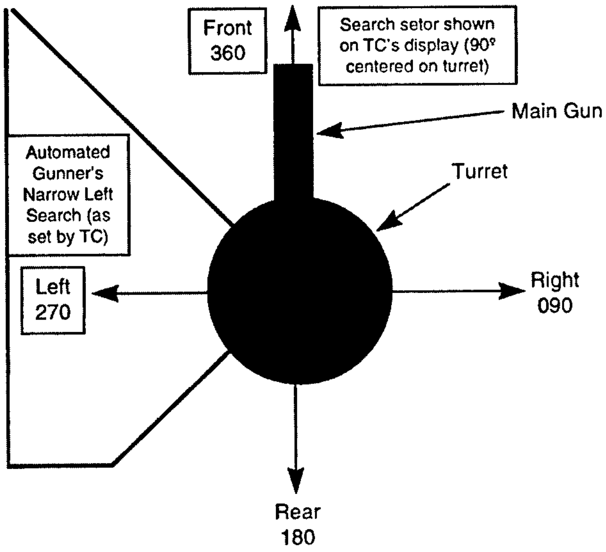


Figure 9. Combined Field of View for TC and automated Gunner's sensors.

Based on analyses like these, this automation concept was adopted as our recommendation, but with minor modifications. It seemed apparent that the TC could take over many of the gunner's tasks given the addition of search sensors and a better method of positioning the gun. Although the task of localizing targets was made somewhat more difficult by the automated gunner's sensors and the Commander's display, overall TC workload was still manageable even in this "worst case" scenario (the target appears in the gunner's sensor area). The TC's increased workload in positioning the gun was comparatively minor, but the overall increase in fire rate was problematic. Since the NTC wanted to emulate human performance, faster-than-normal fire rates were undesirable and we recommended that the automated gunner be delayed approximately 3 seconds to better emulate real human performance. Finally, analysis of the separate workload channels contributing to the final peak in the engagement sequence (that corresponding to repositioning the sensor) showed that this was largely a cognitive problem rather than a visual, manual or verbal one. Relating this to the domain implied that the TC was having problems mentally determining the current and desired position of the automated gunner's sensors. Proposed methods for resolving this problem included slaving the sensors to the hull rather than the turret, and/or including a sensor FOV display in the Commander's display.

5. Facility/Resource Requirements

Once built, our models have proven extremely easy to modify in order to address design or permutation questions. We have used these analytic tools to explore hardware and software changes in proposed cockpits and exploring variations in crew mixture, task allocations and operational procedures. At one point in the advanced attack/scout helicopter design process, we performed 24 permutation analyses, loosely corresponding to 24 crew station redesigns, during a single week.

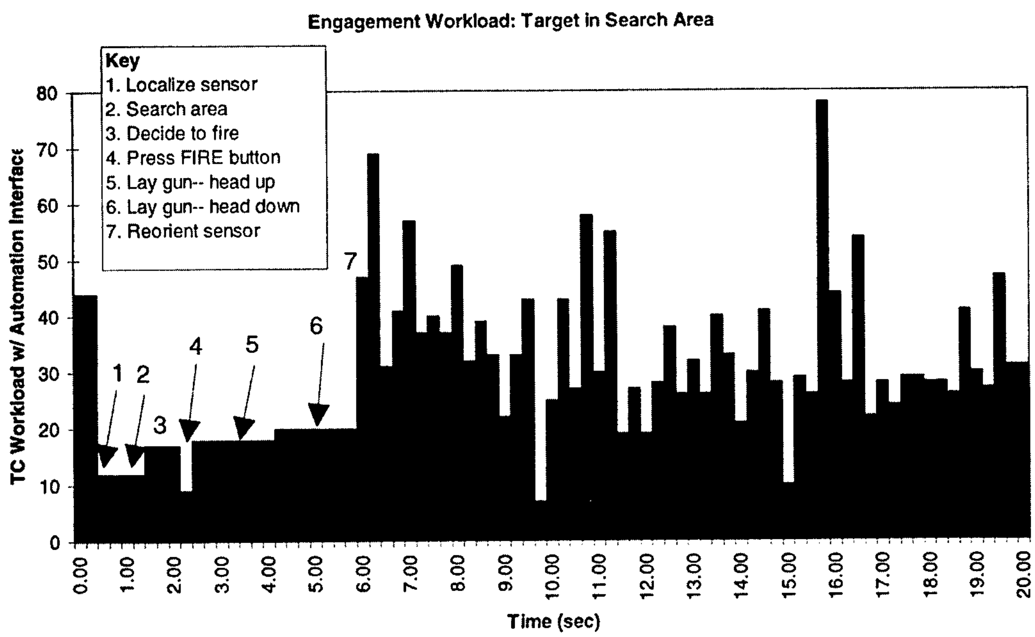


Figure 10. TC's overall workload estimate during engagement scenario in 2 crew condition with sensors and controls as described-- W/Index output.

While the use of workload calculations to evaluate human-crew station interaction is not new, these have generally been used to assess overall operator workload rather than to *predict* specific timesharing conflicts that provide useful data during design. Our approach provides a comparatively inexpensive and rapid method of obtaining useful information about human interaction with a crew station long before even the roughest prototypes are built — information which can be used to focus, refine, and thereby shorten later prototyping and evaluation efforts.

6. ACKNOWLEDGMENTS

The attack/scout helicopter design studies described in this paper were performed under the Rotorcraft Pilot's Associate contract (DAAJ02-92-R-0037); Bruce Tenney and Ray Higgins (AATD/RPA), contract monitors. The NTC study was performed by Tom Plocher as part of the Crew Reduction in Armored Vehicles Ergonomics Study (DAAA15-89-C-0021) with John Lockett of the U.S. Army Human Engineering Laboratory as program monitor. W/Index was developed by Bob North and Victor Riley with Honeywell Internal Research and Development funds. The authors would like to thank Vic Riley and Tom Plocher for their help in performing this work and for comments on earlier drafts of this paper.

7. REFERENCES

1. Wickens, C. "Attention, time-sharing, and workload". In C. Wickens, (Ed.)*Engineering Psychology and Human Performance.* Columbus, OH; Charles E. Merrill, Pub., 1984, pp. 291-334.

2. Bailey, R. W., *Human Performance and Engineering.* (2nd edition). New Jersey: Prentice Hall, 1989.

3. North, R.A. & Riley, V. "W/Index: A predictive model of operator workload". In *Applications of Human Performance Models to Systems Design,* New York; Plenum Press, 1988.

4. Aldrich, T., Szabo, S. and Bierbaum, C. , "The development and application of models to predict operator workload during system design". In G, McMillan et. al. (Eds.) *Applications of Human Performance Models to System Design,* New York; Plenum Press, 1989.

5. Riley, V., Lyall, E., Cooper, B., and Wiener, E. (1993). *Analytic methods for flight-deck automation design and evaluation, Phase One Report: Flight crew workload prediction.* Tech Report for FAA Contract DTFA01-91-C-0039.

Worked Example of the Use of WINCREW in the Evaluation of Overall System Performance

Dr R Laughery Jr
Micro Analysis & Design
4900 Pearl East Circle
Boulder, Colorado 80301
USA

The WinCrew Tutorial

WinCrew is a tool for constructing system performance models for existing or conceptual systems when a central issue is whether the humans and machine will be able to handle the workload. WinCrew can be used to predict operator workload for a crew given a design concept. WinCrew also has the ability to model and predict the effects of that workload on crew and system performance.

What separates WinCrew from other workload models is this direct link between task-induced workload and the effect on system performance. With WinCrew, you can predict how the human will dynamically alter his behaviour when he or she encounters high workload situations. WinCrew can simulate the following as a function of high workload:

- dynamic allocation of tasks between humans, machines,
- dropping tasks based on task priority
- task time and accuracy degradation

The Theory behind WinCrew's Prediction of Human Response to Workload

The best human factors design aid for studying how design and operations concepts will affect the system's performance when human's are being pushed - WinCrew is a human factors tool designed to examine how crew size and design complexity affect mission performance. It provides users with a method to assign workload estimates to tasks that crew members are performing and use those workload estimates to dynamically model the impact on task and system performance. With WinCrew, you can address overall system performance consequences of total crew size and stress as well as the potential value of automation concepts to support high workload scenarios.

WinCrew lets users test theories of how humans manage workload or stress. Users can apply workload management strategies in order to study how the crew will react in times of high workload, and how that reaction will ultimately affect total system performance. Users select from a list of common management strategies including task dumping performance degradation and many others.

WinCrew is based on sound theories of human response to workload. WinCrew implements the Multiple Resource Theory of workload to predict workload. The basis of the workload prediction technique is an assumption that excessive human workload is not usually caused by one particular task required of the operator. Rather, it is the human having to perform several tasks simultaneously that leads to overload, such as drive while they read information off of a display. Since the factors that cause this type of workload are intricately linked to these dynamic aspects of the human's task requirements, task network modeling provides a good basis for studying how task allocation and sequencing can affect operator workload.

However, task network modeling is not inherently a model of human workload. The only relevant output common to all task network models is the time required to perform a set of tasks and the sequence in which the tasks are performed. Time information alone would suffice if workload was to be estimated by comparing the time available to perform a group of tasks to the time required to perform the group of tasks. However, it has long been recognized that this simplistic analysis misses many aspects of the human's tasks that influence both perceived workload as well as ensuing performance. At the very least, this approach misses the fact that some pairs of tasks can be performed in combination better than other pairs of tasks.

The most promising theory of operator workload to emerge over the last 20 years is the multiple resource theory proposed by Wickens (e.g., Wickens, Sandry, and Vidulich, 1983). Simply stated, the multiple resource

theory suggests that humans have not one information processing resource that can only be tapped singly but several different resources that can be tapped simultaneously. Depending upon the nature of the information processing tasks required of a human, these resources would have to process information sequentially (if different tasks require the same types of resources) or possibly in parallel (if different tasks required different types of resources).

WinCrew implements the Wickens' Theory of Multiple Resources. WinCrew supports the hierarchical decomposition of missions into functions and tasks. Tasks are assigned to human resources as well as to the physical interfaces of the workspace. Each task is assigned a workload single task demand value for the resources and interfaces used. For instances when a single operator must execute two tasks at the same time, a workload conflict value is assigned. The WinCrew tool contains a knowledge base of benchmark values for single task demands and channel conflicts. However, users can enter their own task demand and channel conflict values. As the model executes, an overall workload value is calculated using a complex algorithm embedded within WinCrew. This algorithm accounts for the current ongoing tasks' single task demands, and the conflicts between and within resource/interface pairs. From this, users can get a moment by moment estimate of crew workload in several cognitive resource channels during the scenario. WinCrew allows the user to define thresholds for workload values. When workload gets too high (i.e., above the user-defined threshold), the user can define how or if the operator will manage workload. Built in workload management strategies include:

- Dynamic task allocation to other crewmembers
- Dynamic task allocation to the machine
- Dumping an ongoing task
- Not accepting the new task that causes overload to occur
- Delaying an ongoing task and accepting the new task
- Accepting overload with a task time performance penalty
- Accepting overload with a task error rate/accuracy performance penalty

All of these can occur at any time during the simulation and can be driven by the circumstances of the scenario as well as system design and task allocation.

In essence, WinCrew provides a tool for representing Multiple Resource theory on how humans respond to high workload. More details of the above theory and some of the details of implementation can be found below and in the *WinCrew User's Manual.*

Building a Sample Model in WinCrew of a Human Driving an Automobile while Using a Cell Phone

To help you understand how you use WinCrew to model human workload, we have developed a simple model of a human driving an automobile and using a cellular telephone as an example of how some of these WinCrew modeling concepts can be applied to a real situation. In this Appendix, we will briefly describe this model and how it was constructed using the human workload modeling tools embedded within WinCrew.

To review this model description most effectively, you should have a copy of WinCrew and the Phone example that is included with the software to follow along with the text. However, this is not essential.

The Basic Idea behind the Model

Over the past ten years, the use of cellular telephones in automobiles has become very common. Recently, there has been evidence linking the use of a cellular telephone in an automobile to increased probabilities of accidents. The reason can be anticipated as an increase in the driver's workload associated with using a cellular telephone while operating a car. This simple model demonstrates how WinCrew could be used to study this issue.

The Task Network

In this model, we will only simulate two functions performed by the driver, driving and talking on the telephone. When they are done with both, the simulation is completed. Therefore, the highest-level model structure includes the functions represented on Figure 1.

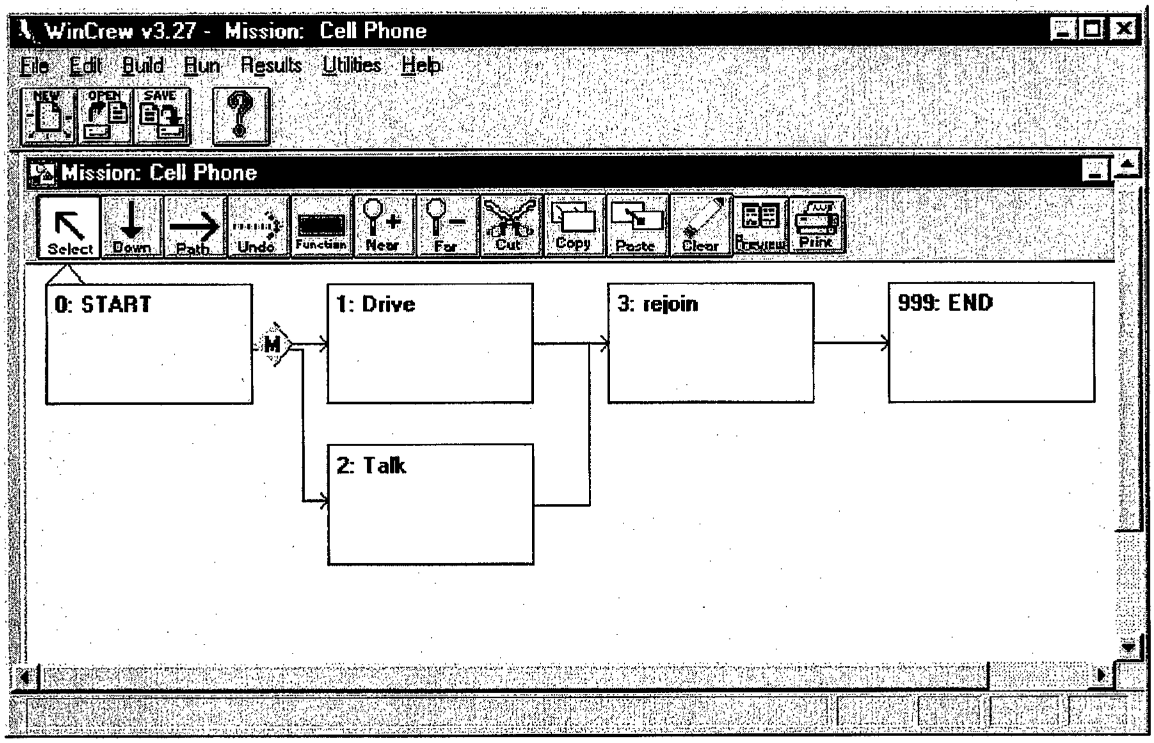


Figure. 1 Functions in the Cell Phone Model

Three of these functions, START, rejoin, and END, do not involve human activity but are required to manage the flow of simulation activities.

The Drive function is modeled as including the tasks as indicated in Figure 2.

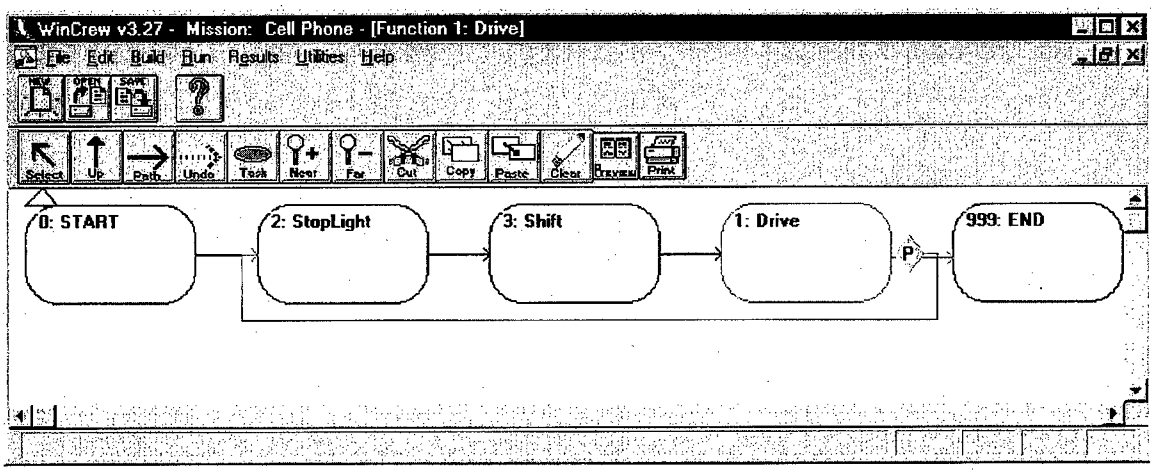


Figure 2. Tasks in the Drive Function

As shown, the simulation begins with the driver sitting at a stoplight, accelerating, and driving until either the simulation ends or another stoplight is approached. In this model, the completion of the model is determined by the probabilistic branch at the end of task 1 as is shown in Figure 3. In a more complex model, the simulation could proceed for a fixed number of stoplights or for a fixed time simply by incorporating the appropriate decision logic at the decision point marked by the "P" after task 1.

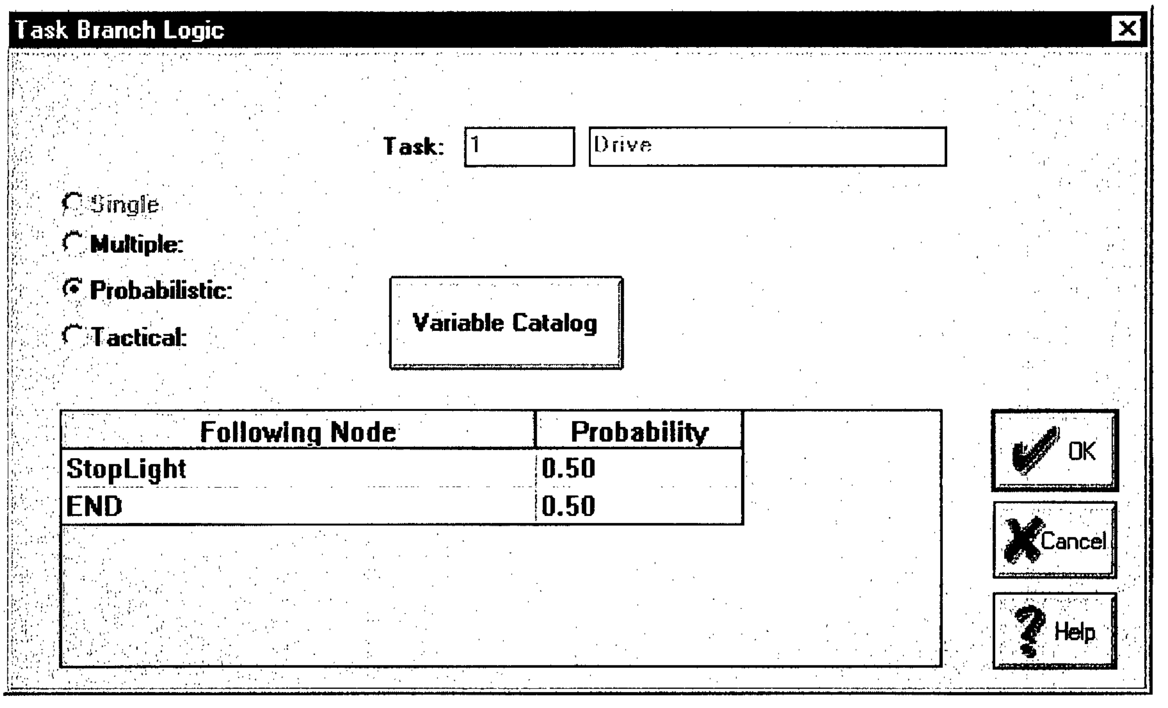
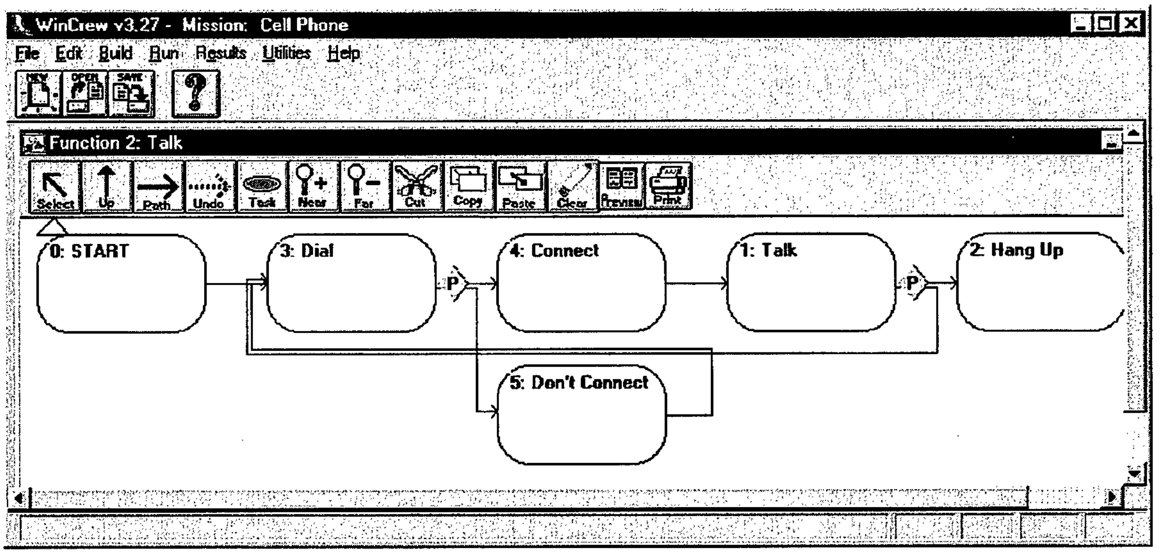


Figure 3. Probabilistic Branch defining likelihood of ending the simulation or approaching another stop light

The task network for the function Talk is presented in Figure 4. It also uses a probabilistic branching approach to simulating the number of telephone calls made by the driver. There is also a probabilistic branch after the Dial task that simulates that some calls do not go through and, therefore, must be redialed.



Each of the tasks in this simulation takes time that is estimated based on existing data. Figure 5 shows the Task Description window obtained by opening up the Dial task.

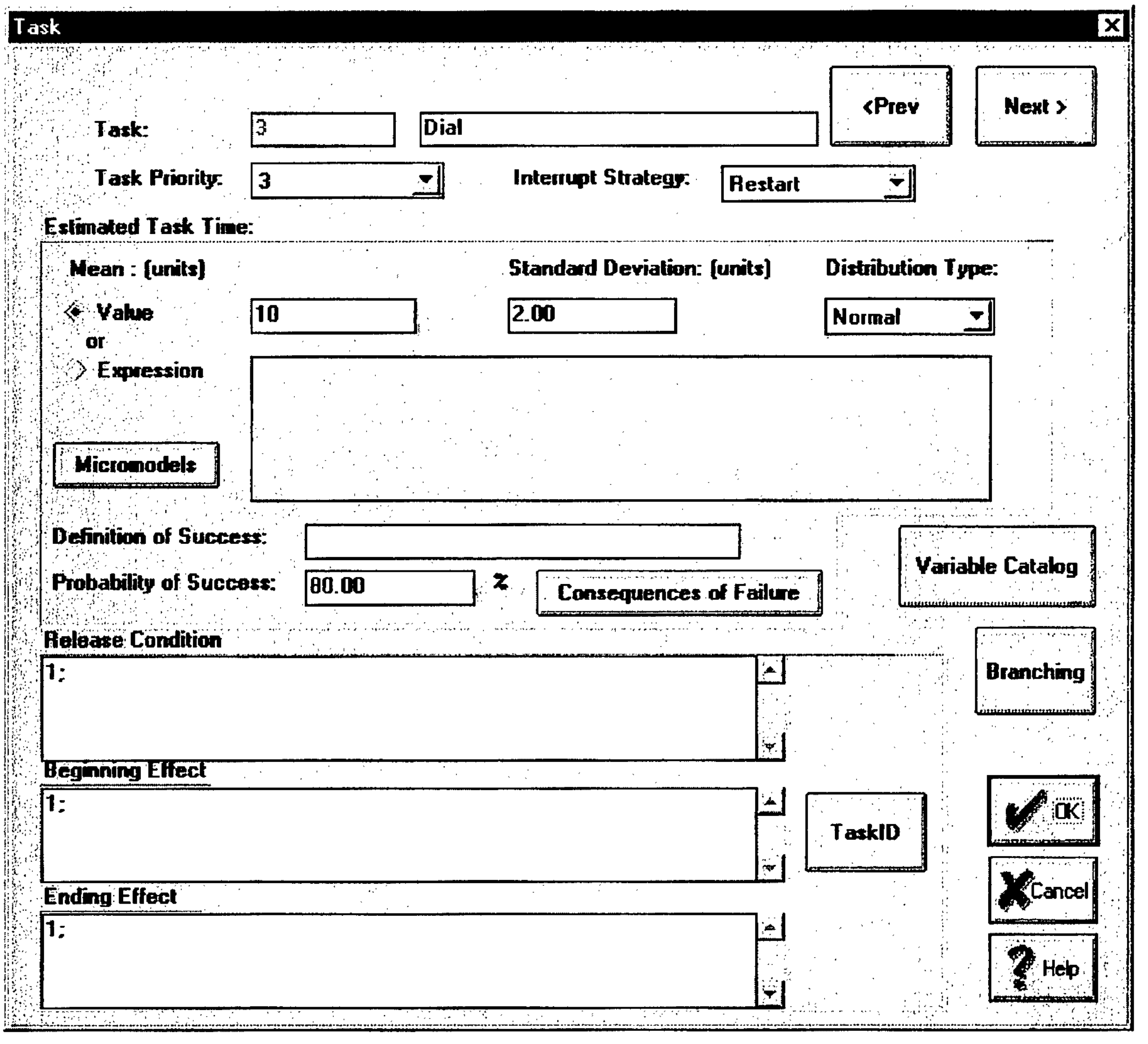


Figure 5. Task Description Window for the Dial Task

This task will take a normally distributed amount of time with a mean of 10.0 seconds and a standard deviation of 2.0 seconds. In this task, no micromodels are used and there are no release conditions required for the commencement of this task and this task has no effects on system parameters when it begins or ends. More complex models may use these fields, but they are not necessary in this simulation.

Also, as shown on Figure 6, there is an 80% likelihood that this task will succeed every time it is performed and, therefore, a 20% chance that it will fail. This simulates, for example, the entry of an incorrect number when entering the telephone number. By selecting the Consequences of Failure button, a window as shown in Figure 7 is opened.

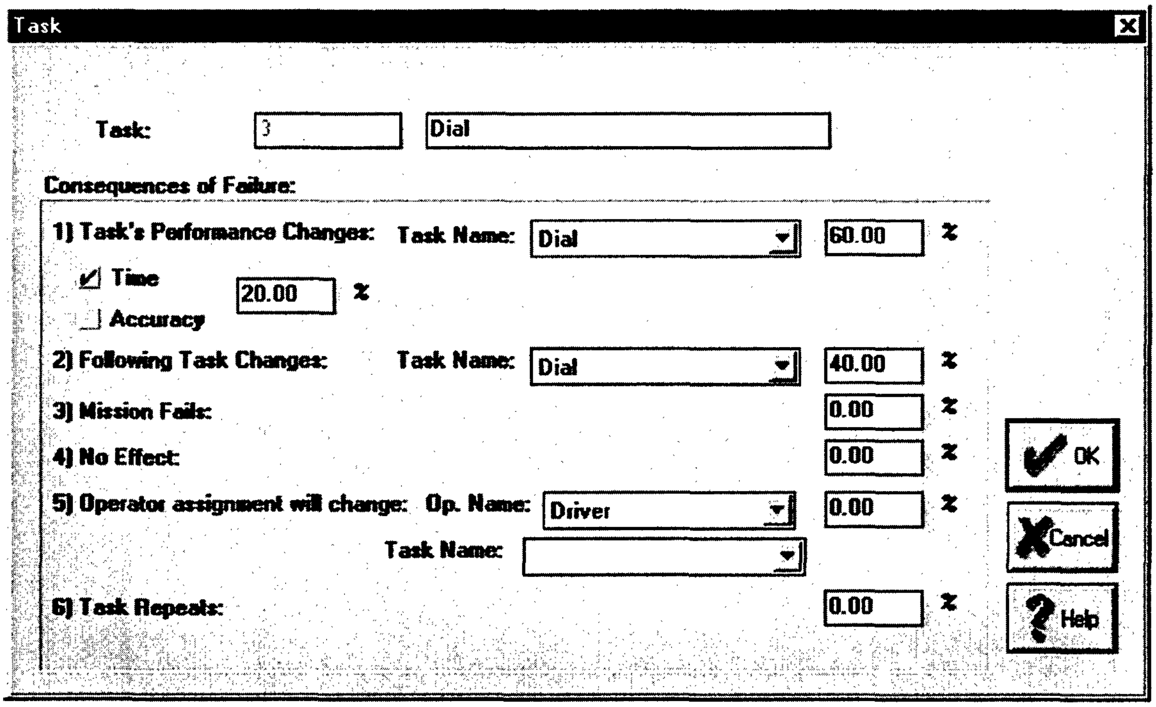


Figure 6. Defining the Consequences of Failing to Dial Correctly

As simulated here, whenever the task fails two possible things might happen. 60% of the time, the time of the Dial task is increased by 20%, representing the time to backspace over the incorrect number and re-enter that number. The other 40% of the time, the whole task will need to be repeated representing the situation where the driver does not notice until they actually finish the whole dialing process.

The probabilistic branch shown after the Dial task in Figure 4 represents whether the connection was made upon completion of the dialing (e.g., if the number was busy or the phone was out of range of a cell). This is also represented by a probability as shown in Figure 7.

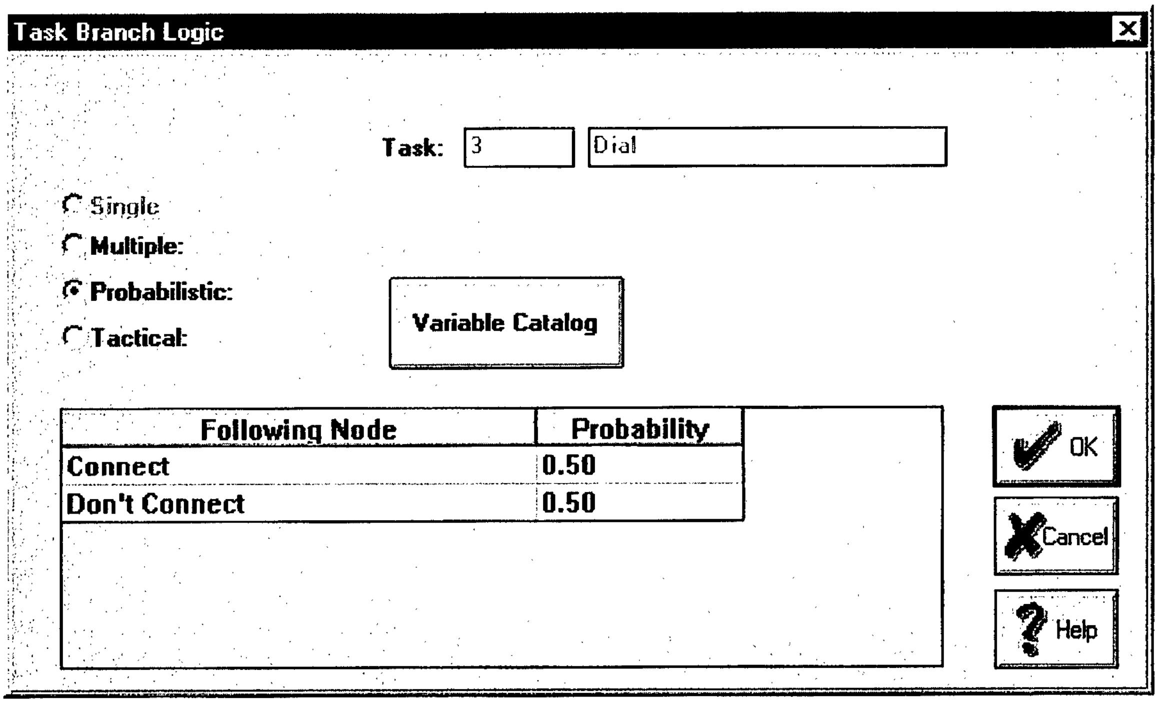


Figure 7. Defining the Probability of Achieving a Connection Once a Number is Dialed

As simulated in this model, the driver will continually attempt to redial the number until a connection is achieved. Once a connection is achieved, the driver will talk for a period of time as represent in the mean time and standard deviation in the task description window for this task as shown in Figure 8.

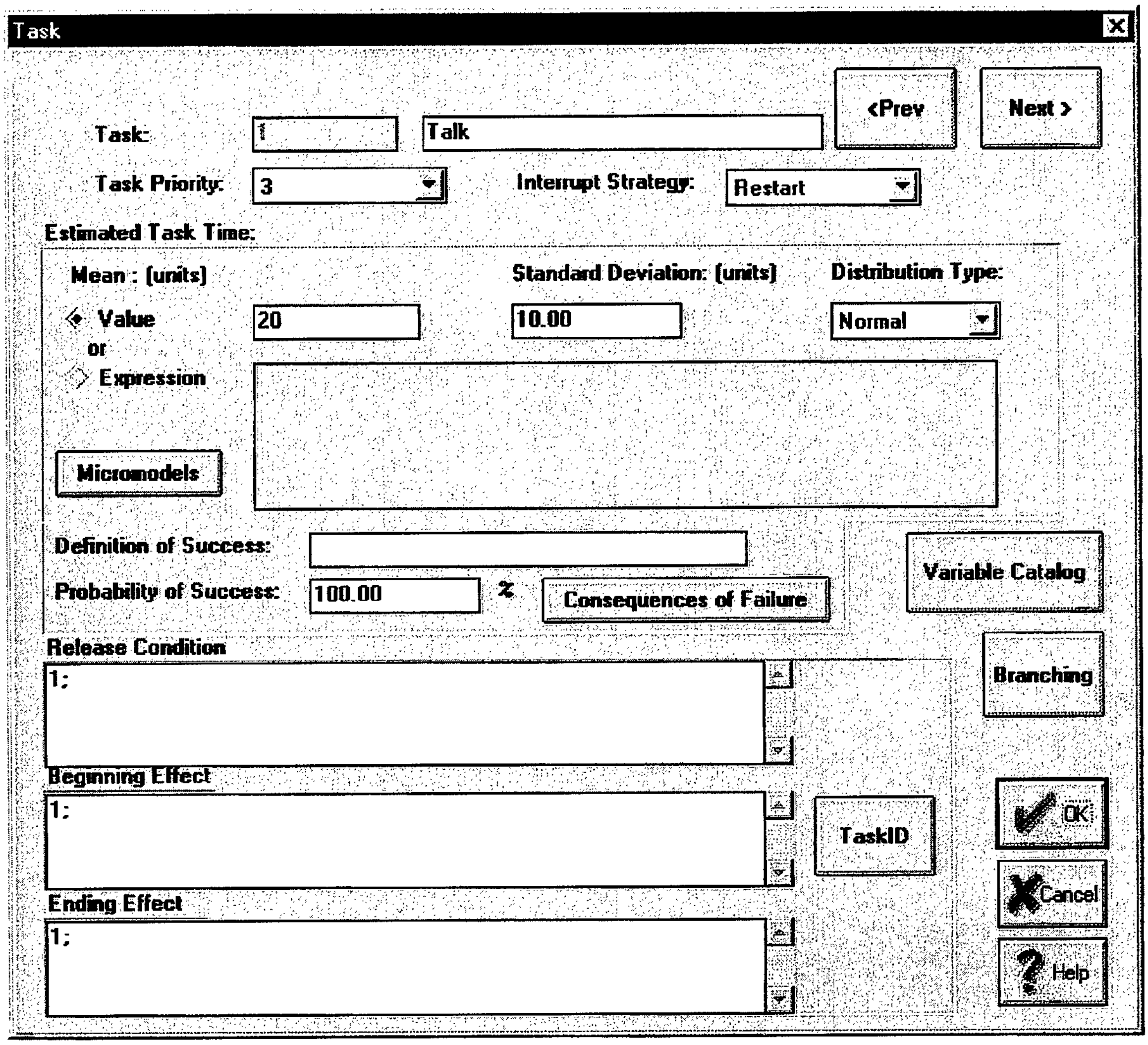


Figure 8. Task Description Window Representing the Task of Talking

You will note that this task has a high standard deviation relative to the mean representing the high variability of telephone call times.

As shown in Figure 4, after the driver is done with a call, there is a probability that another call is made. If not, the use of the cell phone is completed. In this model, the probability that another call will be made is 75%.

Defining the Operators, Task Assignments, and how High Workload will be Managed

In this model, we are simulating only one operator. To define an operator, we select the *Define Operators* menu option, which is a sub-menu off of the *Crewmembers and Automation* option off of the *Build* menu. Figure 9 presents the Define Operators interface with the information filled out for the driver. If other options are selected later on (e.g., an inexperienced or a fatigued driver), then simulated performance of the driver will be modified as described in the WinCrew manual.

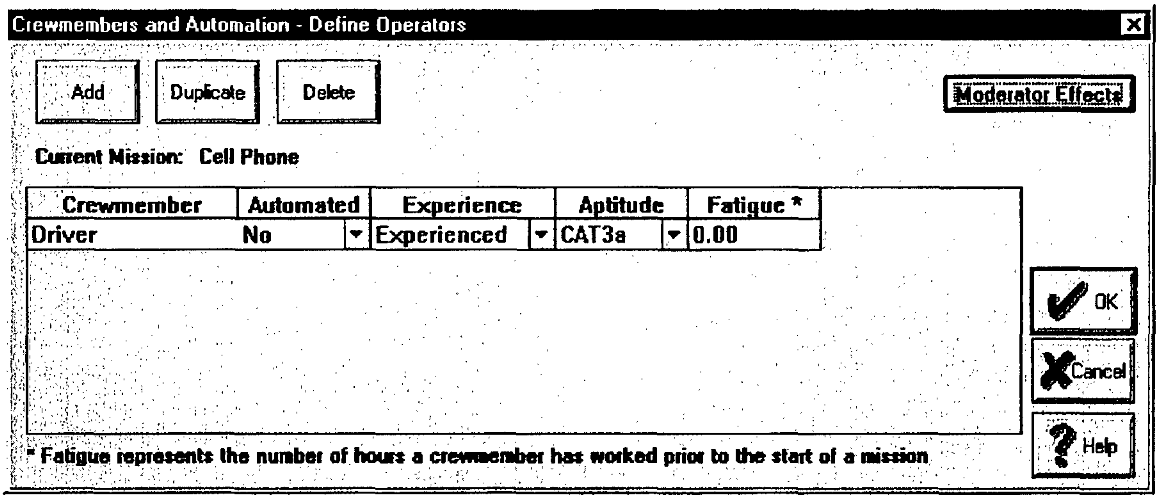


Figure 9. Define Operators Interface for Defining a Driver

Under the *Task Assignment* interface as shown in Figure 10, the primary and contingency operators are defined. In this model, all tasks are assigned to the primary operator. However, if we wanted to simulate the potential assistance that a passenger might provide, we could define an operator called *Passenger* and then assign some of the telephone tasks to the Passenger as a contingency operator to perform when workload gets too high on the driver.

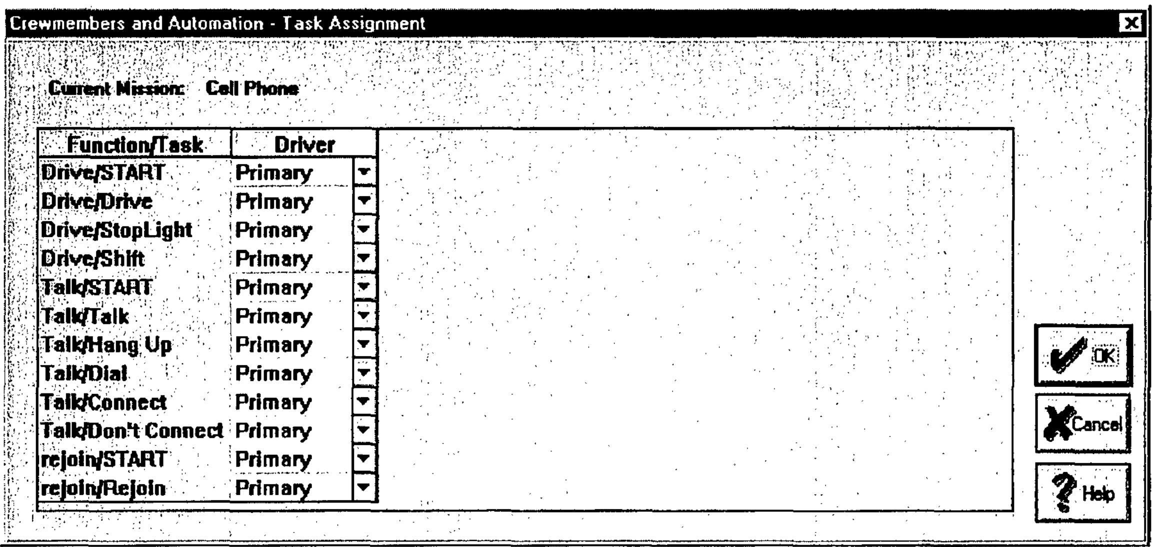


Figure 10. Task Assignment for the Driver

The next step will be to define *Workload Management* on the interface as shown below which is also available from the *Crewmembers and Automation* sub-menu under the *Build* menu. Workload management refers to what the operator will do when a new task that the operator is scheduled to begin will place the operator beyond the workload threshold. The value of the threshold that will force the operator to go into workload management is also defined in this interface. The workload management is defined for this model in Figure 11.

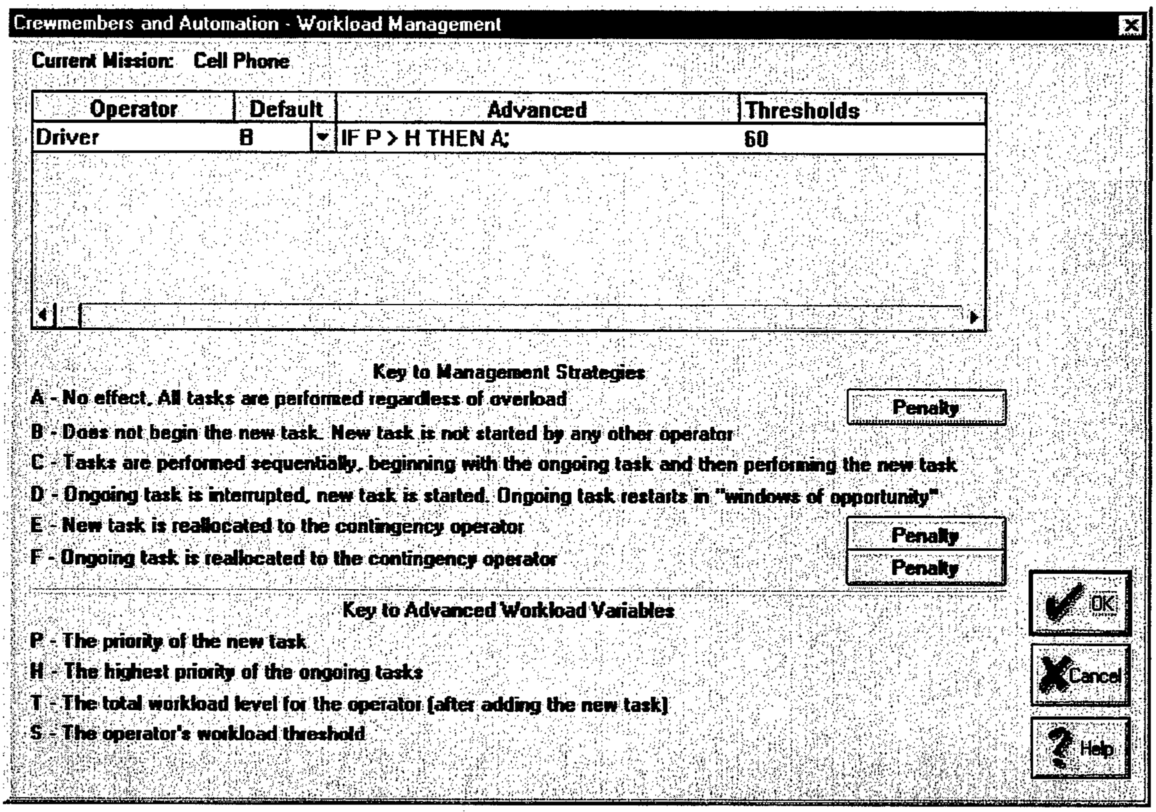


Figure 11. Defining workload management

As defined in this interface the driver will go into an overload situation whenever the new task will cause the workload value to exceed a value of 60. The default management strategy when this occurs is management strategy A which, as defined in the Key to Management Strategies portion of the screen, is that the driver will accept the new task and, in essence, nothing will change. If we chose to define a penalty associated with this strategy, we could simply press the Penalty button to the right of the description and define the Penalty in terms of either a task time increase or an increase in the probability of an error. However, *if* the new task's priority is less than the priority of any of the ongoing tasks, then the management strategy adopted will be Strategy B, or that the driver will not accept the new task.

In this model, we have defined the priority of the driving tasks to be higher than the tasks associated with the telephone. Therefore, the effect of this strategy is that a driving task will always be performed, even if it forces the driver into high workload. However, if dealing with the telephone will force the driver into high workload, the driver will not perform the telephone task and all use of the phone will stop.

Defining the Operator Interface and How It Drives Workload

To estimate workload, we must define the interface elements and the workload attached to using them in various tasks. All of these are defined from the *Workload and Crewstation Parameters* sub-menu, which is off of the *Build* menu

You begin this by selecting the Resources and Interfaces sub-menu. For this model, the resources and interfaces that are defined are shown in Figure 12.

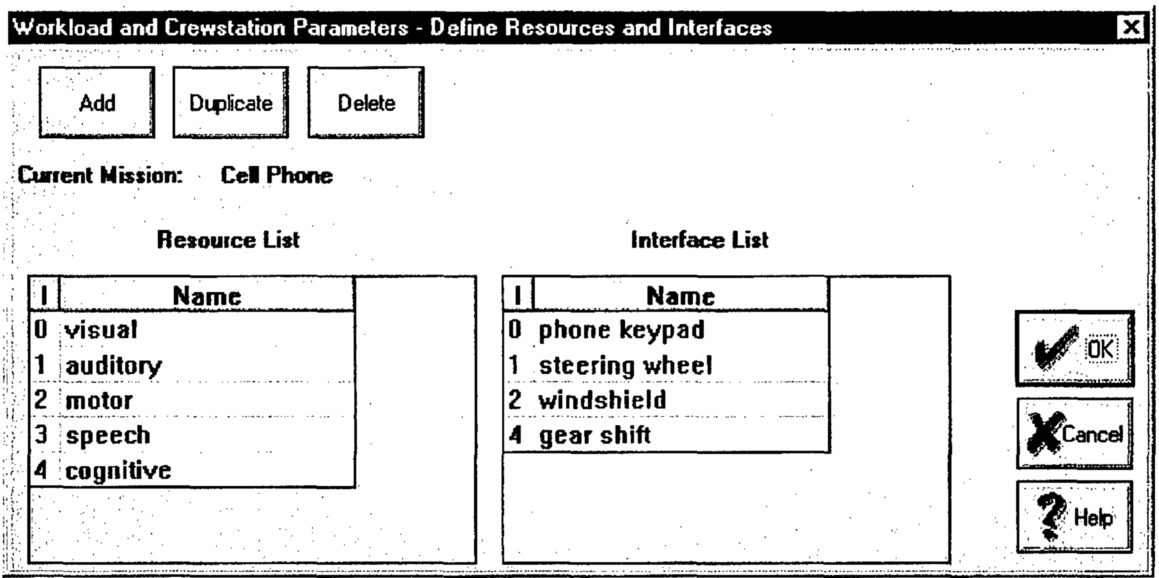


Figure 12. Resources and Interfaces

The resource list shown in Figure 12 is the standard list that comes with WinCrew. The four interfaces shown were entered by the modeler.

Next, the resource/interface channel combinations need to be defined. These define the resources that are required for interacting with each interface. Figure 13 presents this interface for this model.

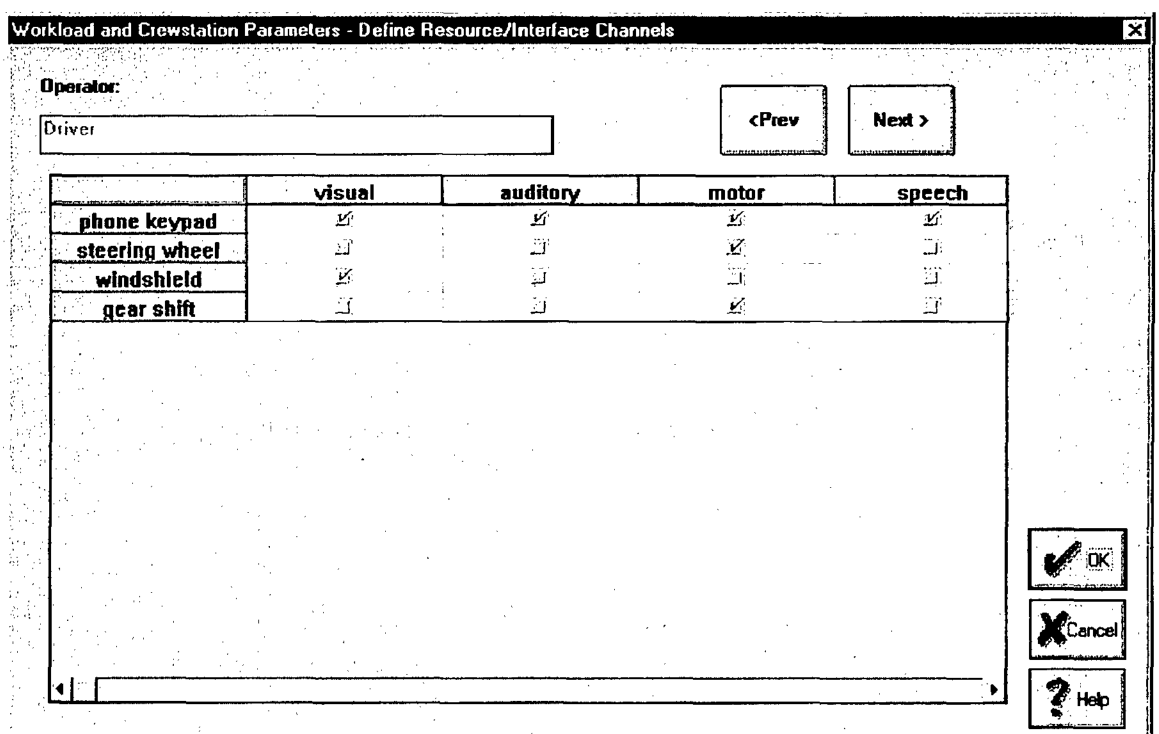


Figure 13. Defining Resource/Interface Channels

For example, from this interface you can see that the windshield requires only visual resources, the gear shift and steering wheel require only motor resources, but the phone keypad requires visual, auditory, motor, and speech resources. Actually, all defined interfaces require cognitive resources as well as would be seen by sliding the viewing bar at the bottom of the screen to the right. The

definition of resource/interface channels is made simply by clicking in the appropriate box with the mouse. If there were multiple operators, the channels would need to be defined for each operator.

Next, the resource interface channels defined above need to be associated with tasks that require those resource interface channels using the interface as shown in Figure 14.

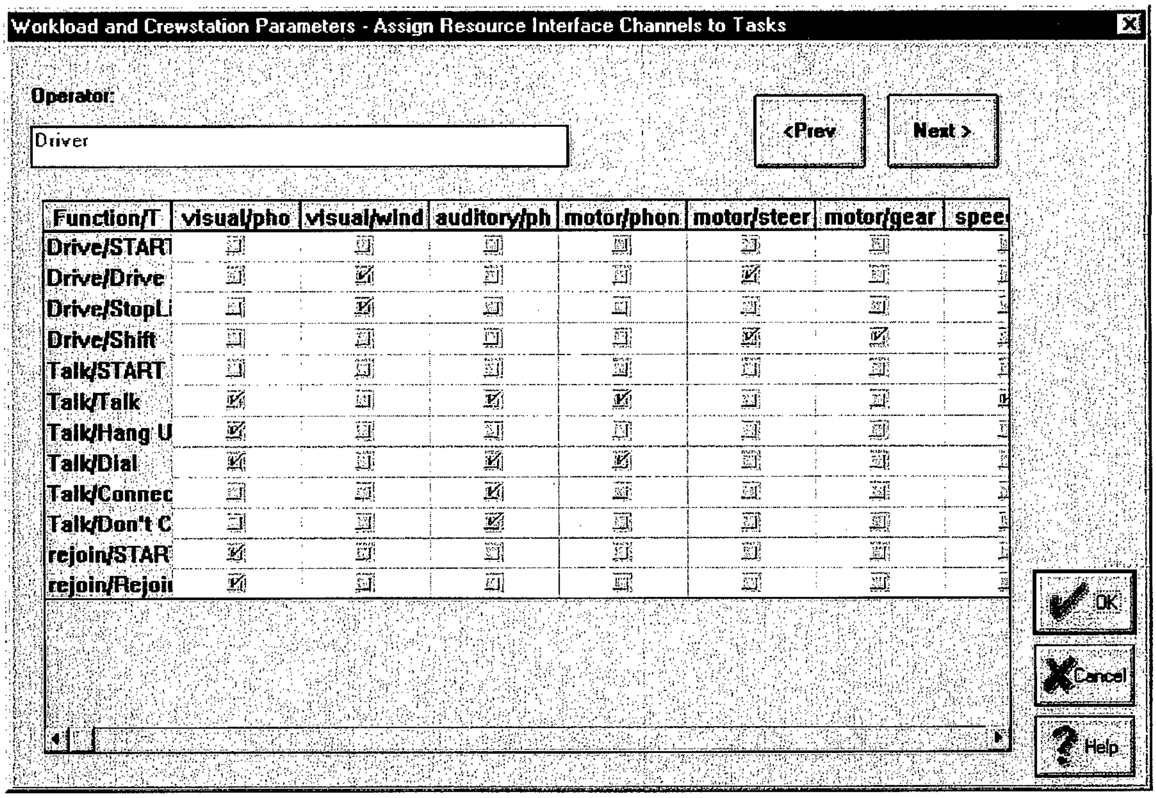


Figure 14. Associating Resource Interface Channels with Tasks

Each task included in the model is listed as a row and each resource interface pair is listed as a column. The resource interface pairs that are used for each task are defined by clicking in the checkbox. Again, if there were other operators, these would be defined uniquely for each operator.

Also, the single task demand values for each resource interface pair must be defined as shown in Figure 15.

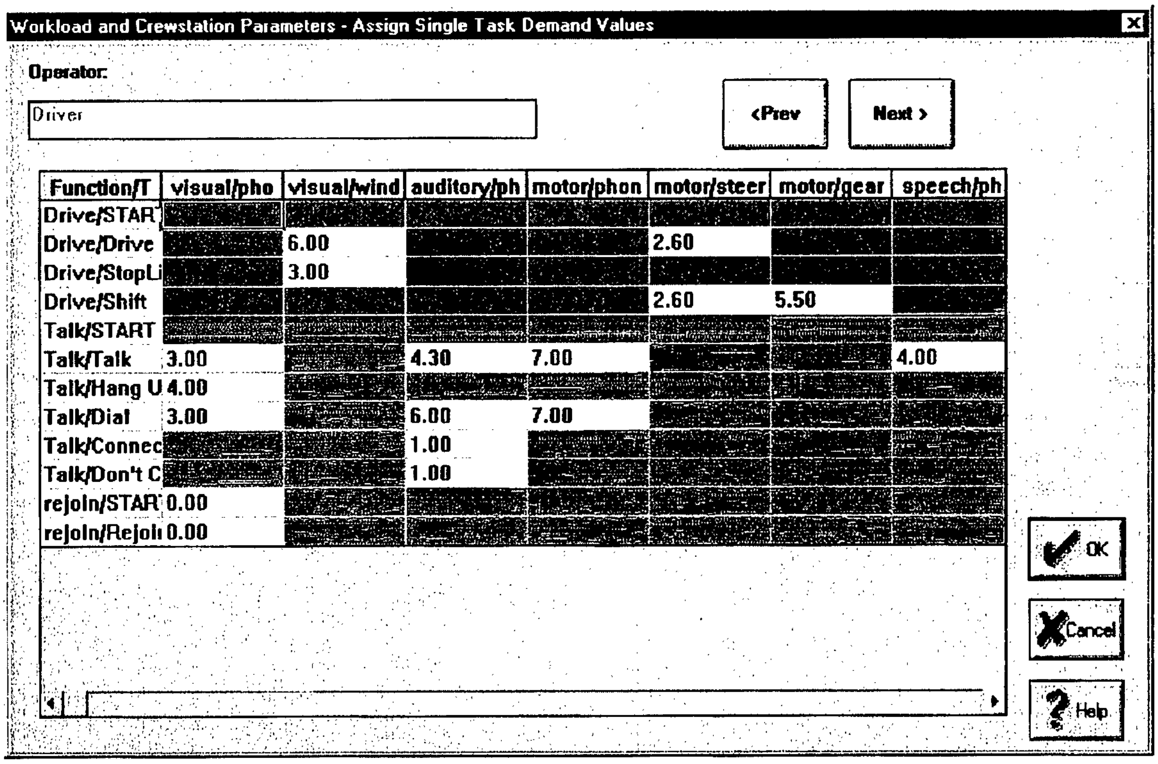


Figure 15. Defining Resource Interface Single Task Demand Values

These values are defined either by entering a value in the cell or by double clicking in any cell that is white (indicating that a resource interface pair has been defined for that task) which will pop up a menu similar to that shown in Figure 16. Different menu options will be presented for different resource categories.

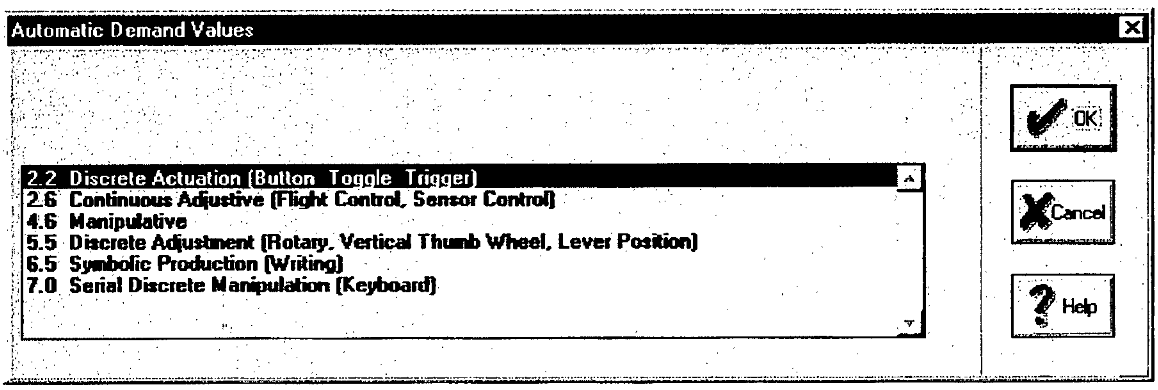


Figure 16. Defining Demand Values Pop up Menu

Finally, to define workload, the channel conflict values must be defined as shown in Figure 17.

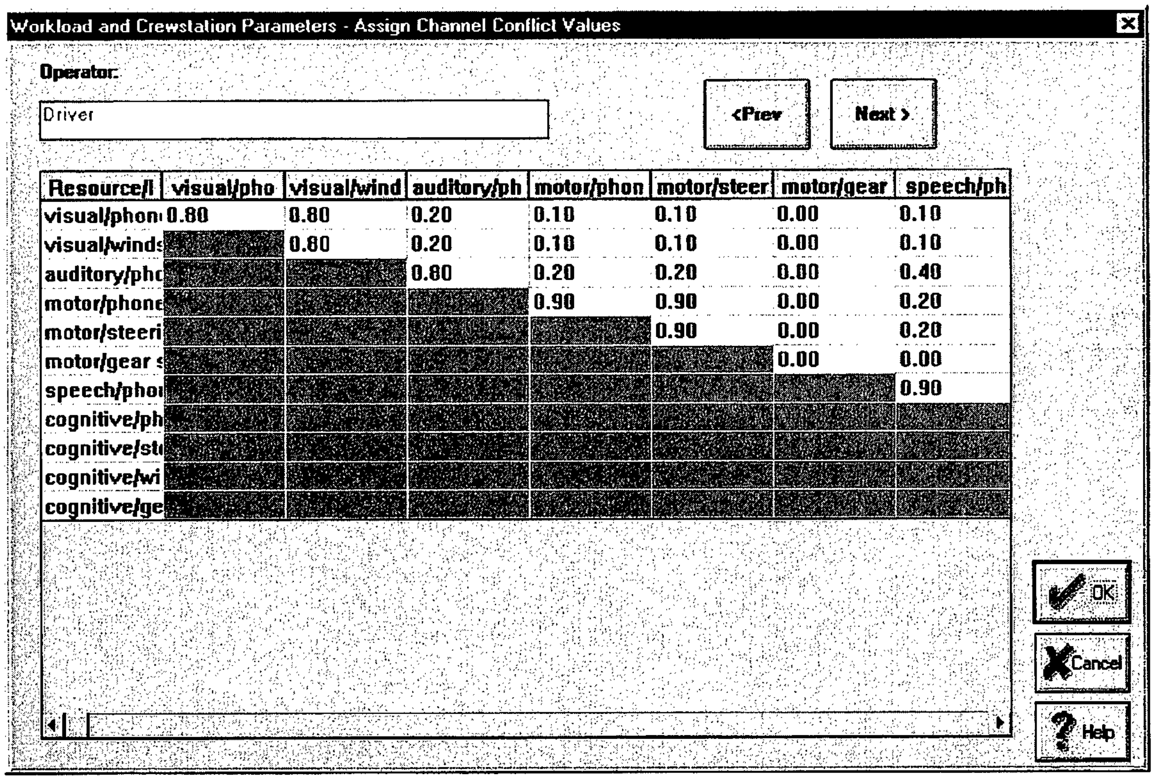


Figure 17. Assigning Channel Conflict Values

These values define the inherent conflicts in trying to perform multiple tasks simultaneously that demanded resource interface pair combinations. For example, it would be very difficult to engage in two motor tasks involving the phone keypad at the same time. Therefore, in the matrix in Figure 17 where the "motor/phone keypad" row and column intersects, a value of 0.9 was entered in the matrix indicating high conflict when this resource interface pair is demanded twice at the same time. Alternately, performing tasks that involve both visual tasks with the windshield and motor tasks with the gear shift involve no inherent conflict, so a value of 0 was entered in this cell.

By defining all of the above, a model of a driver using a cell phone has been built in WinCrew.

Executing the Model and Reviewing Results

To run the model, select *Execute Model* from the *Run* menu. A pop-up menu as shown in Figure 18 will appear.

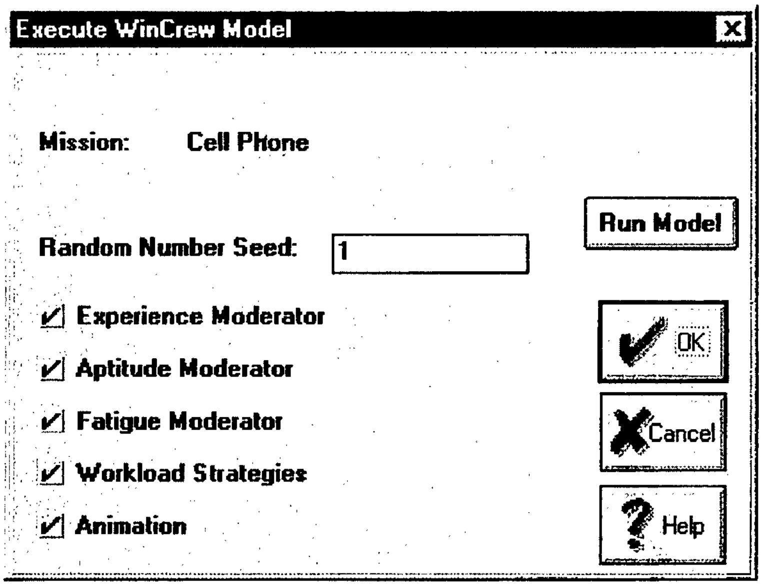


Figure 18. Model Execution Options

In this menu, the user can select whether to use WinCrew built-in algorithms to modify task time and accuracy associated with experience, aptitude, and fatigue. Also, the user can turn on or off workload strategies. By turning this off, the model will not simulate modifying operator behavior in high workload situations using workload management strategies. Selecting animation will allow a display of the task network as it runs with animation. Animation involves highlighting tasks as they are executing as shown in Figure 19.

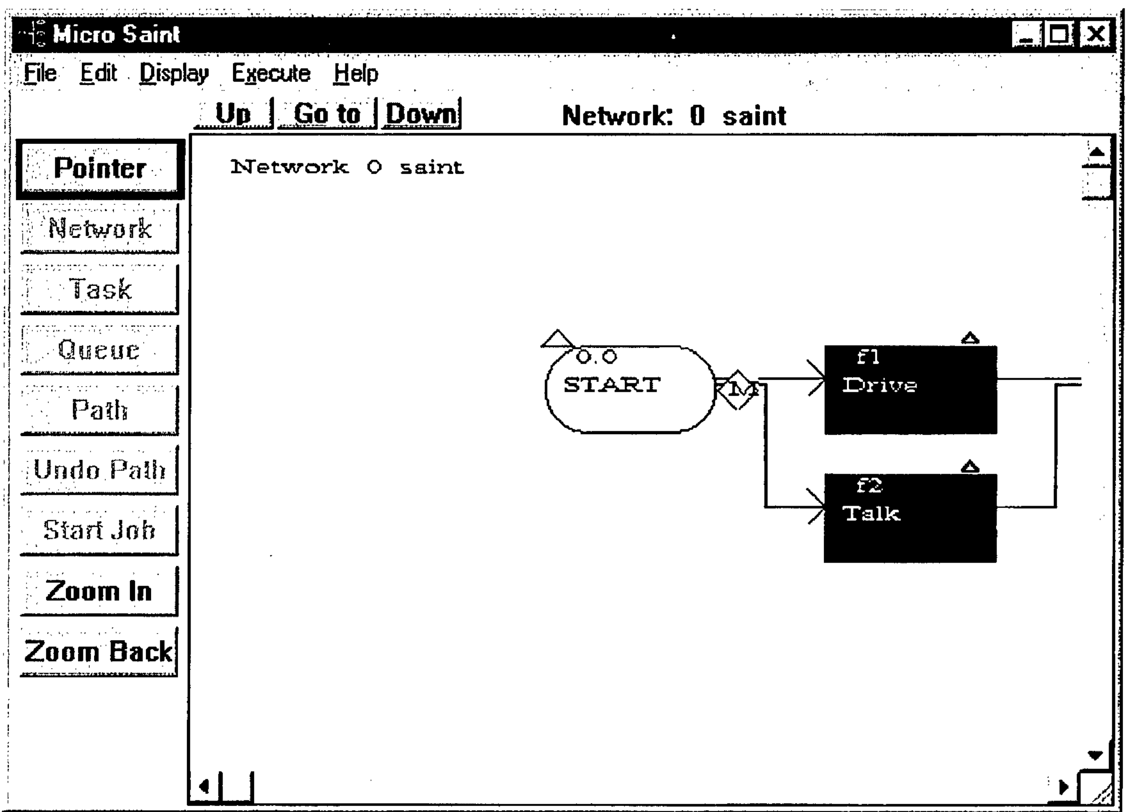


Figure 19. Model Animation Interface

In complex models, animation can be helpful in determining tasks that must often be performed simultaneously. In a relatively simple model such as this, it may not be needed.

The types of reports available to the user are shown in Figure 20.

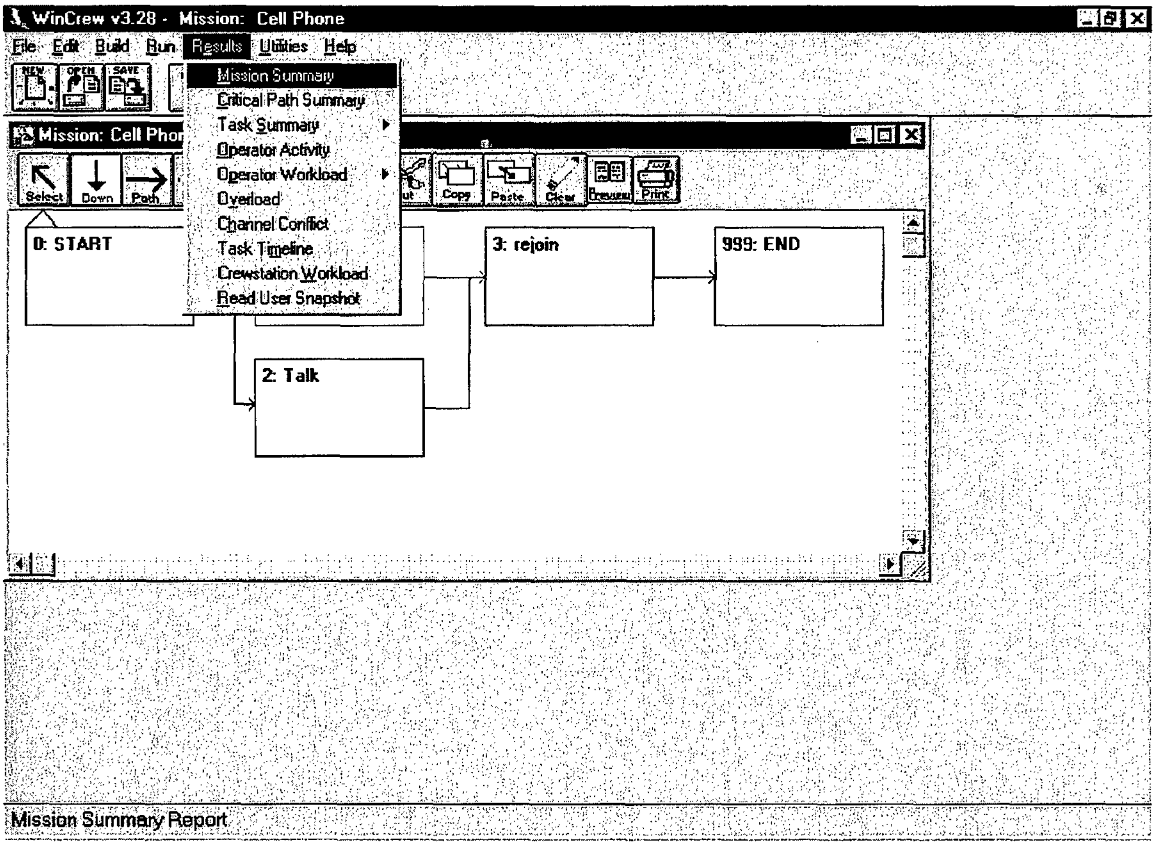
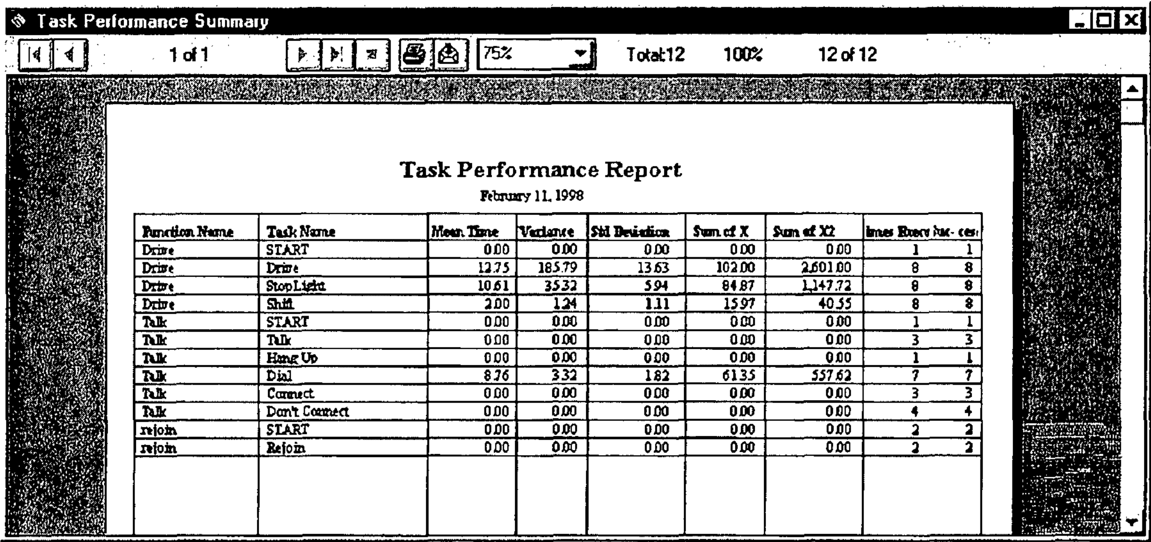


Figure 20. List of Reports Available from Results Menu

For this model, the interesting reports are the Task Summary, Operator Activity, Operator Workload, Overload, Channel Conflict, and Task Timeline. All or portions of each of these reports are presented in Figures 21 through 26, respectively.

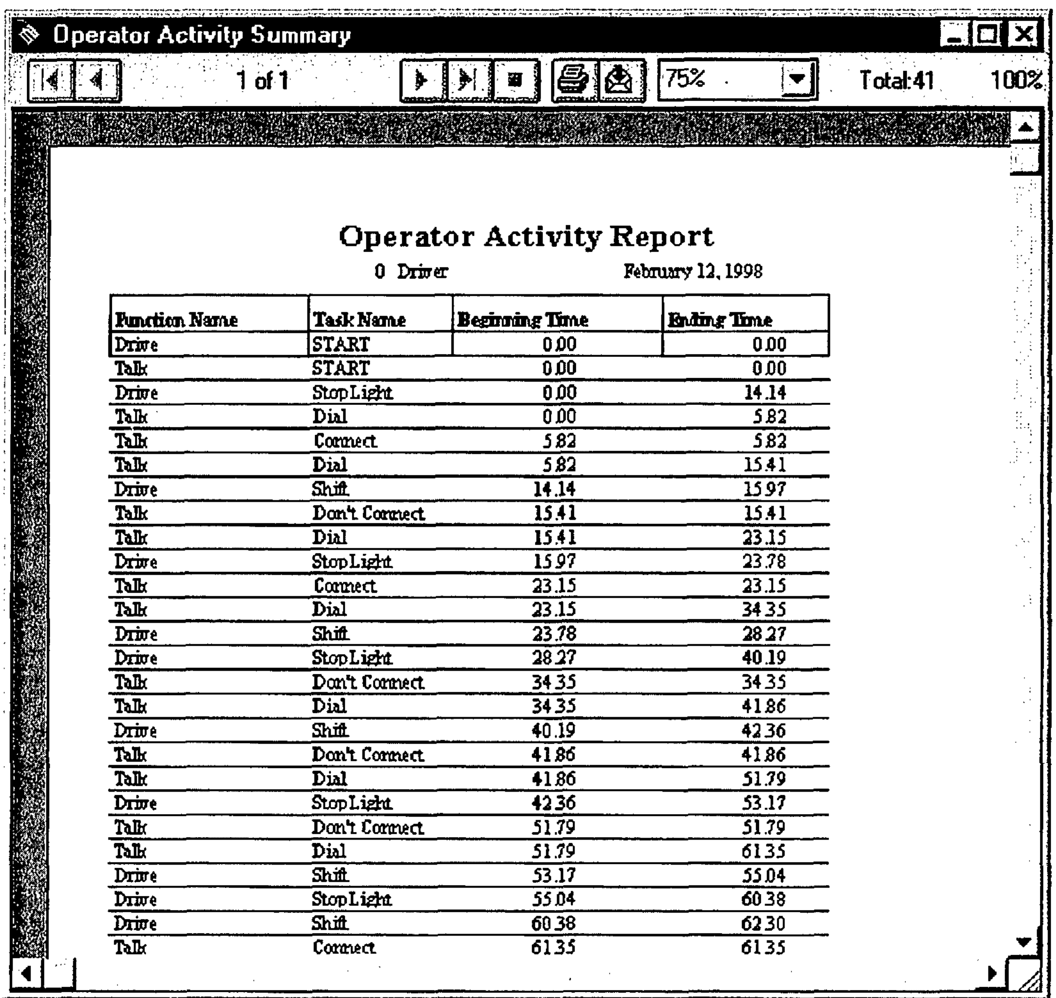


Task Performance Report

February 11, 1998

Function Name	Task Name	Mean Time	Variance	Std Deviation	Sum of X	Sum of X2	Times Execu	Suc- ces
Drive	START	0.00	0.00	0.00	0.00	0.00	1	1
Drive	Drive	12.75	185.79	13.63	102.00	2,601.00	8	8
Drive	Stop Light	10.61	35.32	5.94	84.87	1,147.72	8	8
Drive	Shift	2.00	1.24	1.11	15.97	40.55	8	8
Talk	START	0.00	0.00	0.00	0.00	0.00	1	1
Talk	Talk	0.00	0.00	0.00	0.00	0.00	3	3
Talk	Hang Up	0.00	0.00	0.00	0.00	0.00	1	1
Talk	Dial	8.76	3.32	1.82	61.35	557.62	7	7
Talk	Connect	0.00	0.00	0.00	0.00	0.00	3	3
Talk	Don't Connect	0.00	0.00	0.00	0.00	0.00	4	4
rejoin	START	0.00	0.00	0.00	0.00	0.00	2	2
rejoin	Rejoin	0.00	0.00	0.00	0.00	0.00	2	2

Figure 21. Task Summary Report



Operator Activity Report

0 Driver February 12, 1998

Function Name	Task Name	Beginning Time	Ending Time
Drive	START	0.00	0.00
Talk	START	0.00	0.00
Drive	Stop Light	0.00	14.14
Talk	Dial	0.00	5.82
Talk	Connect	5.82	5.82
Talk	Dial	5.82	15.41
Drive	Shift	14.14	15.97
Talk	Don't Connect	15.41	15.41
Talk	Dial	15.41	23.15
Drive	Stop Light	15.97	23.78
Talk	Connect	23.15	23.15
Talk	Dial	23.15	34.35
Drive	Shift	23.78	28.27
Drive	Stop Light	28.27	40.19
Talk	Don't Connect	34.35	34.35
Talk	Dial	34.35	41.86
Drive	Shift	40.19	42.36
Talk	Don't Connect	41.86	41.86
Talk	Dial	41.86	51.79
Drive	Stop Light	42.36	53.17
Talk	Don't Connect	51.79	51.79
Talk	Dial	51.79	61.35
Drive	Shift	53.17	55.04
Drive	Stop Light	55.04	60.38
Drive	Shift	60.38	62.30
Talk	Connect	61.35	61.35

Figure 22. Operator Activity Report

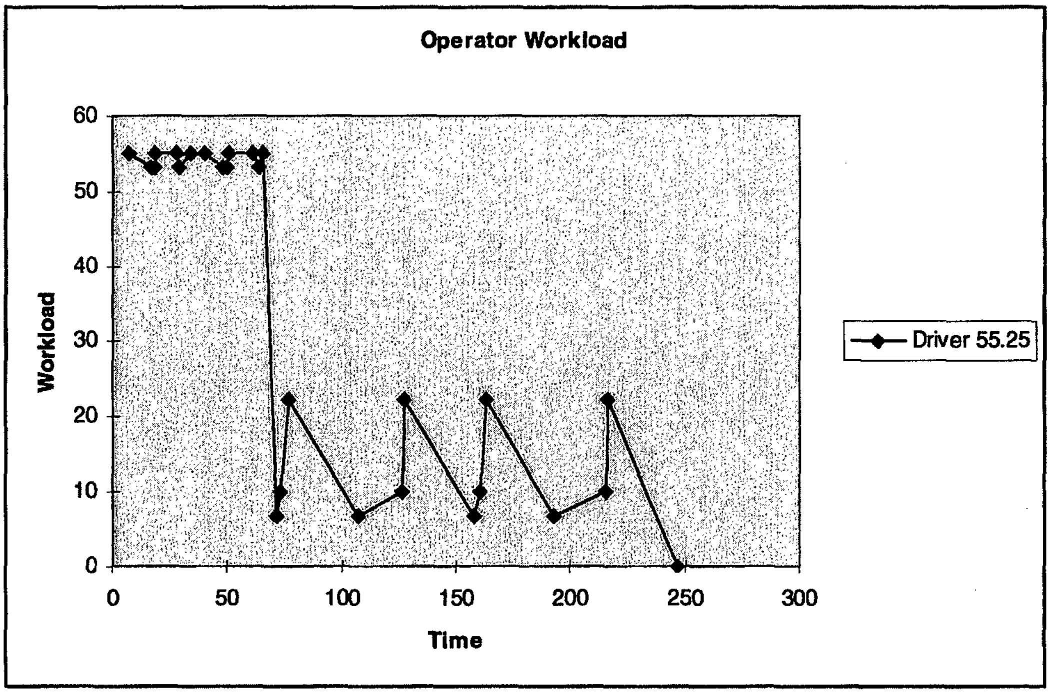


Figure 23. Operator Workload Report

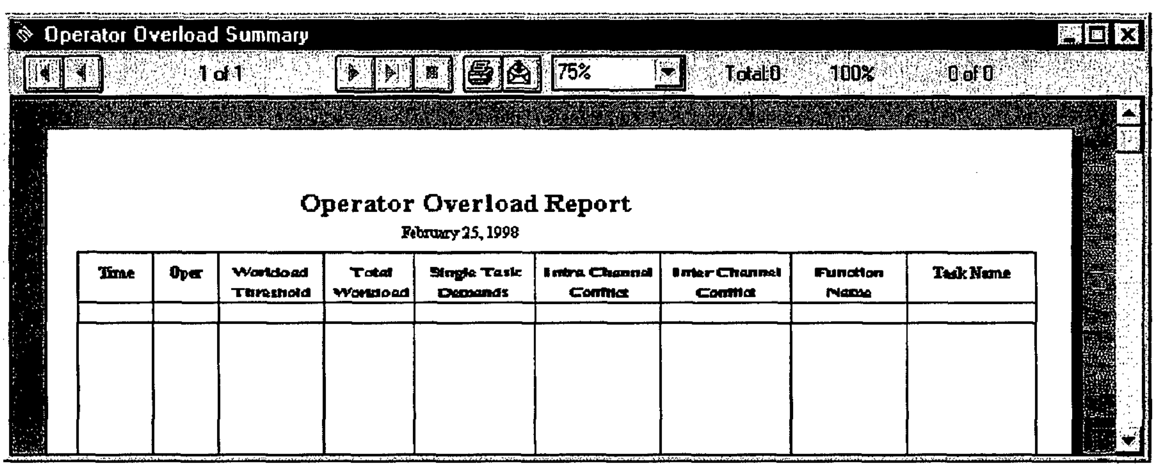


Figure 24. Overload Report

Channel Conflict

1 of 1 75% Total:45 100% 45 of 45

Channel Conflict Summary

February 12, 1998

Operator: **0 Driver**

Resource/Interface Pair1	Resource/Interface Pair2	Conflict Value	# of Times Exercised
visual/phone ke	visual/phone ke	.8	7
visual/windshie	visual/phone ke	.8	3
auditory/phone	visual/phone ke	.2	0
motor/phone key	visual/phone ke	.1	0
motor/steering	visual/phone ke	.1	3
speech/phone ke	visual/phone ke	.1	0
cognitive/phone	visual/phone ke	.5	0
cognitive/steer	visual/phone ke	.5	3
cognitive/winds	visual/phone ke	.5	3
visual/windshie	visual/windshie	.8	7
auditory/phone	visual/windshie	.2	4
motor/phone key	visual/windshie	.1	3
motor/steering	visual/windshie	.1	0
speech/phone ke	visual/windshie	.1	4
cognitive/phone	visual/windshie	.5	4
cognitive/steer	visual/windshie	.5	0
cognitive/winds	visual/windshie	.5	0
auditory/phone	auditory/phone	.8	8
motor/phone key	auditory/phone	.2	0
motor/steering	auditory/phone	.2	3
speech/phone ke	auditory/phone	.4	0
cognitive/phone	auditory/phone	.4	0
cognitive/steer	auditory/phone	.4	3
cognitive/winds	auditory/phone	.4	4
motor/phone key	motor/phone key	.9	7
motor/steering	motor/phone key	.9	3
speech/phone ke	motor/phone key	.2	0
cognitive/phone	motor/phone key	.1	0
cognitive/steer	motor/phone key	.1	3
cognitive/winds	motor/phone key	.1	3

Figure 25. Channel Conflict Report

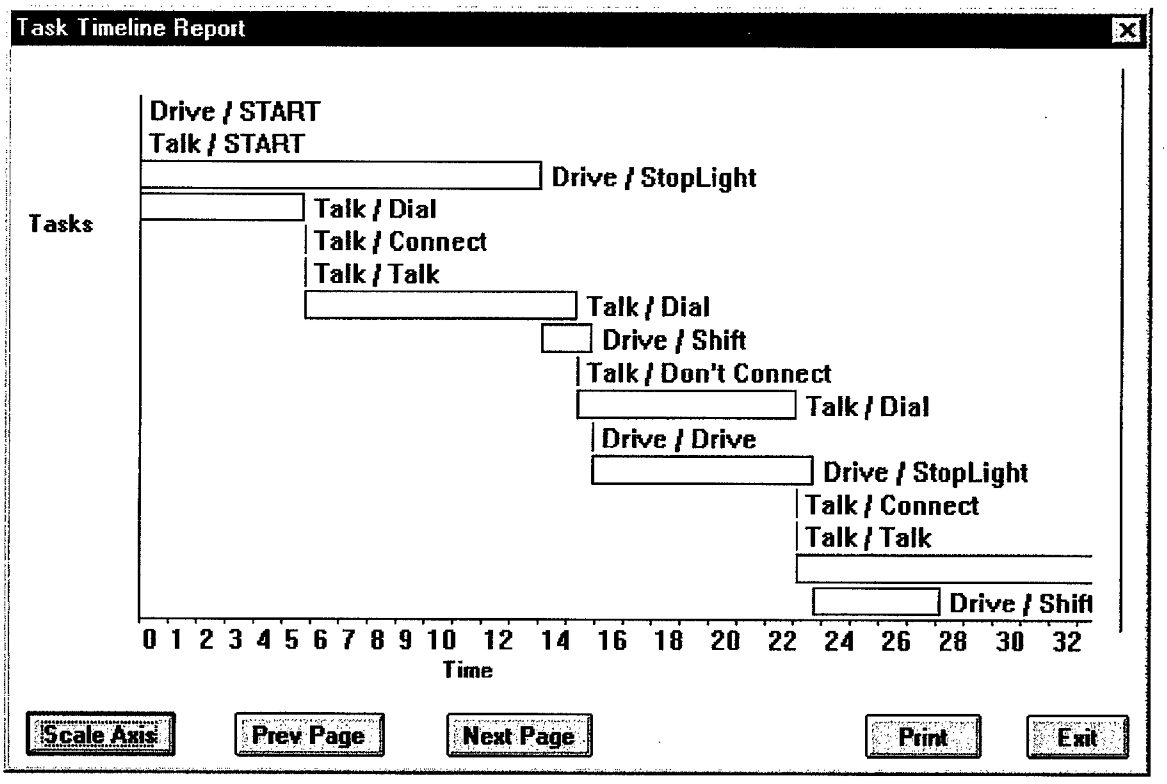


Figure 26. Task Timeline Report

As can be seen from a review of the above reports, there is clearly an effect of the use of a cell phone on workload, although not to the point where the driver is driven in to overload. However, in this simple model, we do not account for difficult driving conditions, unexpected other events that might occur and demand attention, or other distractions that are sometimes present like a radio or another person. These other more pressing situations could be modeled in WinCrew and the effect of using a car phone could be studied simply by making additions to the above model.

Summary

The above very simple WinCrew model illustrates many of the key features that make WinCrew useful for studying system design, task allocation, and task management strategies on system performance. While the above model is fairly small and simple, it captures the elements of behavior that cause many systems to become at risk because of high operator workload.

Man-Machine Integrated Design and Analysis System (MIDAS) FUNCTIONAL OVERVIEW

Kevin Corker
Christian Neukom
NASA Ames Research Center
Moffett Field, Ca. 94035-1000

The following series of screen print-outs illustrates the structure and function of the MIDAS system. Views into the use of the system and editors are featured. The use-case in this set of graphs includes the development of a simulation scenario ....

SLIDE 1: “TOP-LEVEL ELEMENTS” : The main software subelements of the MIDAS system are illustrated here.. The user enters the system through the Graphical User Interface (GUI) that provides the main interaction between the designer and the MIDAS system. The user selects among four functions in the system. Generally the sequence would require the user to establish (create and/or edit) a domain model (which includes establishment and selection of the parameters of performance for the human operator model(s) in the simulation. The user can then select the graphical animation or view to support that simulation or a set of simulations. The user can specify in the simulation module the parameters of execution and display for a given simulation set, and specify in the results analysis system the data to-be-collected and analyzed as a result of running the simulation. The results analysis system also provides for archival processes for various simulation sessions.

The user would typically use all of the top-level features to support a new simulation. If a user were exploring, for instance, the assignment of function between a human operator and a automated assistant the user could maintain the majority of the extant domain, graphical and analytic models and make modification through the domain model to the human operator model, to the equipment model and to the simulation scenario.

SLIDE 2: “RECAP MILESTONE 1: DOMAIN MODEL”: The domain model consists of descriptors and libraries supporting the creation of:

• Vehicle characteristics- (location space, aerodynamic models of arbitrarily detailed fidelity, and guidance models for vehicle (automatic) control.

• Environment characteristics- including terrain form selected data bases at varied levels of resolution, weather features in so far as they effect vehicle performance or operator sensory performance, and cultural features (towns, towers, wires etc.) In short, the analyst here specifies the world of action of the experiment/simulation.

• Crew-Station/Equipment characteristics- the crew station design module and library is a critical component in the MIDAS operation. Descriptions of discrete and continuous control operation of the equipment simulations are provided at several levels of functional detail. The system can provide discrete equipment operation in a stimulus-response (black-box) format, in a time-scripted/event driven format, or in a full discrete space model of the transition among equipment states. Similarly the simulated operator’s knowledge of the system can be at the same varied levels of representation, or can be systematically modified to simulate various states of misunderstanding the equipment function.

• The Human Operator Model(HO)- the human performance model in MIDAS allow for the production of behavior and response for single and multiple operators in the scenarios. The human operator model is the key to the MIDAS function as a predictive design aid. The HO is composed of integrated functions as submodels which include an anthropometric model, sensation and perception models, attention (and other resource models), central processing cognitive functions such as decision making, evaluation and action selection, and finally behavioral models to guide the anthropometric model in the execution of action.

• Mission and Activity Models: Describe in a hierarchic structure the goals and the available recovery activities from missions-not-as-planned that make up the human operators high level behavioral repertoire in the mission. The next level of decomposition of the action of the mission is a set of high level procedures (that can be stored as a fairly generic set of routines, e.g. look-at or fixate). Finally there are the specific actives in "active action packets" RAPS that are the process by which the human operator affects the simulation.

SLIDE 3 "CREWSTATION EDITOR: Illustrates the editing tool of the crewstation domain model with three different access modes, outline, structure and geometry views Modification to the crew station equipment are undertaken in this editor with function and geometry (CAD packages) available for modification.

SLIDE 4: HUMAN PERFORMANCE MODEL: OVERVIEW: The human operator performance model is a combination of a series of functionally integrated micro-models of specific cognitive capabilities within a human operator. The human operator model functions as a closed-loop control model with inputs coming from the world and action being taken in the world. The model provides psychological plausibility in the cognitive constructs of long-term, working memories (with articulation into spatial and verbal components of the theses models) and with sensory/perceptual and attentional components that focus, identify and filter simulation world information for the operator, action and control. The cognitive function is provided by the interaction of context and action. Context is a combination of declarative memory structures and incoming world information is mapped to the agenda manager which is taking the plan (overall mission). This combined with with the plan interpreter provide a series of RAPS to be performed in order to meet mission goals and to handle contingent activities (like interruption or plan repair). Output of action in the world is effected through the models of the operator linked to the anthropometric representations (if they are invoked by the analyst). The action changes the external world and the cycle begins again.

SLIDE 5: VISUAL MODE: EXTERIOR SCAN: Illustrates a process of visual acquisition of external information. The timeline at the bottom illustrates the time for the physical and perceptual components of the scan process and the column on the left illustrates a "situational awareness function" that has been recently developed for the MIDAS system (Shively and Goodman, 1998). The information form the visual scan moves trough states of processing and awareness as more information is made available to the cognitive processor. The data on which situation is based moves form physical information (Detected) to more abstract semantic data found in the long term memory declarative information centers of the operator (recognized) to the final assignment of a definitive identification. These cognitive activities (as with most actual cognitive activities) take time and effort to perform.

SLIDE 5: RAP REVIEW: Provides a detailed look at the sketchy plan operation of the reactive action packet (RAP) implementation of the MIDAS activity structure (Firby 1998) The RAP consists of a set of methods that interpret the context of the current set of goals relative to the sketchy plan and selection action to move the simulation to the desired state.

SLIDE 6: TASK AGENDA: The agenda structure stores instantiated RAPS as goals with subnetworks and logical control flags, object bindings and history of state and completion. This network represents the current set of tasks to be performed by the operators of the simulation given the current goals and context. The network can complete successfully, be interrupted by other task networks or be aborted. The relationship among the actions in terms of logic of performance (e.g. sequential or concurrent tasks) is also specified in the agenda structure. Whether in fact tasks can be performed concurrently is a function of resource relations in the cognitive model (sensation/reception, central/attentional/effectors))

SLIDE 7: PROVISIONING: The provisioning system is the underlying framework for managing the input data for a MIDAS simulation. Input data includes model, scenario, and simulation parameter specifications. The provisioning system provides for fully dynamic specification of the scenario, flexible access to model and simulation libraries, and input data specification.

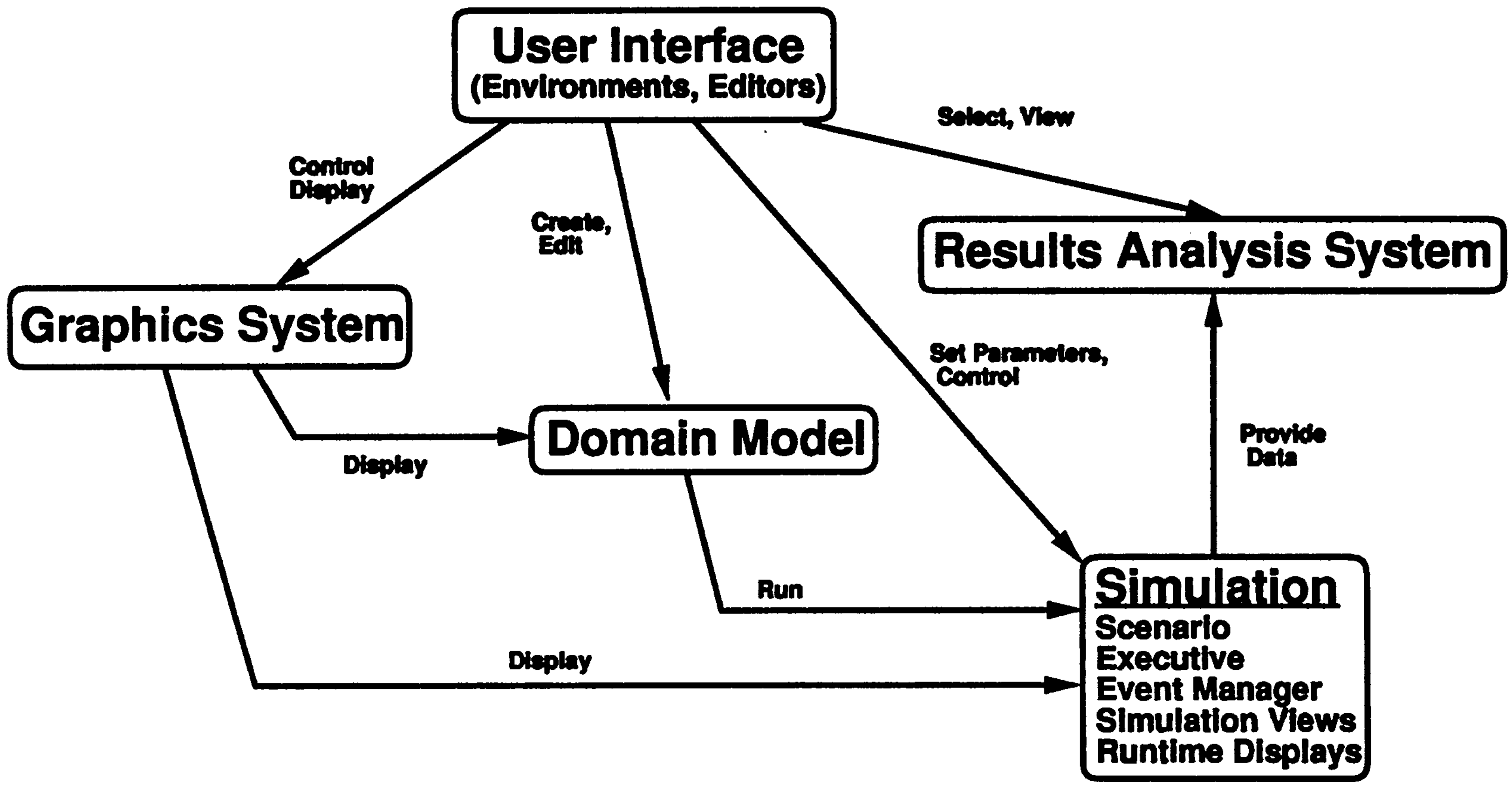
SLIDE 8: DATA ANALYSIS OVERVIEW: Provides a view to the typical kinds of analyses and examinations that can be undertaken in the simulation data runs. Task time history, loads on resource-limited channels, and links for any time to the chain of simulation events is a commonly required feature. More elaborate statistical analyses in a post hoc fashion comparing the time-histories of one run versus another are also available

SLIDE 9 THOUGH 11 : BASIC SCENARIO. These represent a series of charts to illustrate a basic operational scenario in which the pilot flies the mission and co-pilot maps the terrain.

Firby, R.J. (1989). *Adaptive Execution in Complex Dynamic Worlds*. Tech Report YALEU/CSD/RR #672, Yale University (Ph.D. Dissertation).

MIDAS Redesign Milestone 3

Top–Level Elements



BACK NEXT

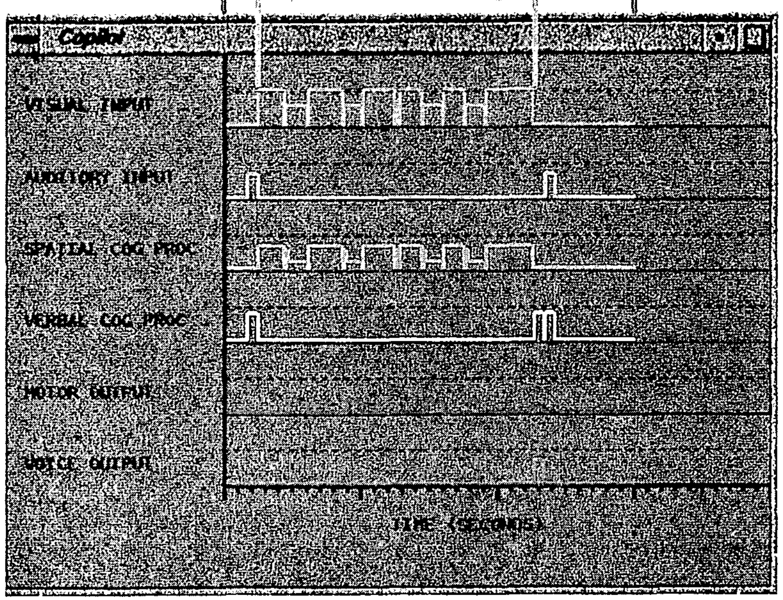
MIDAS Redesign Milestone 3

Data Analysis Overview

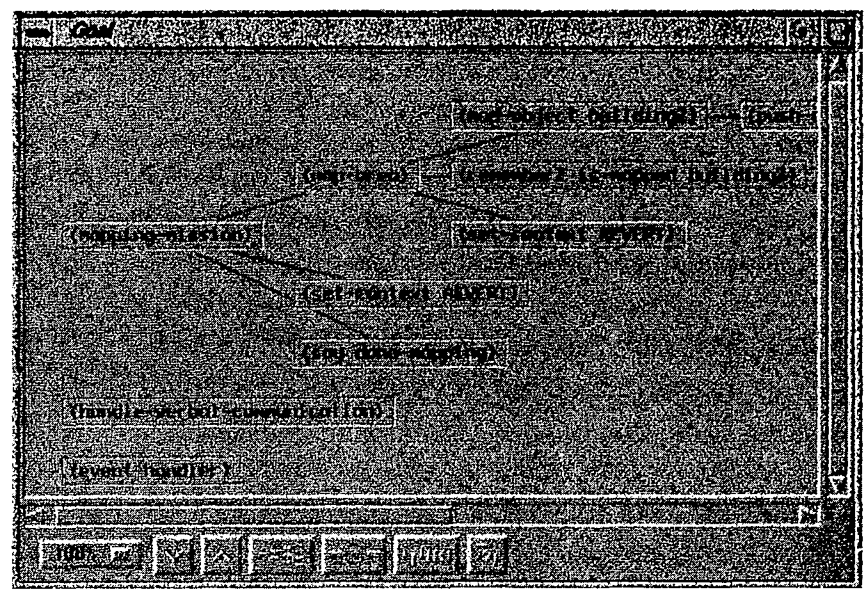
Data analysis can do the descriptive statistics for attention demands

Overall data analysis

Task based analysis (mapping mission)

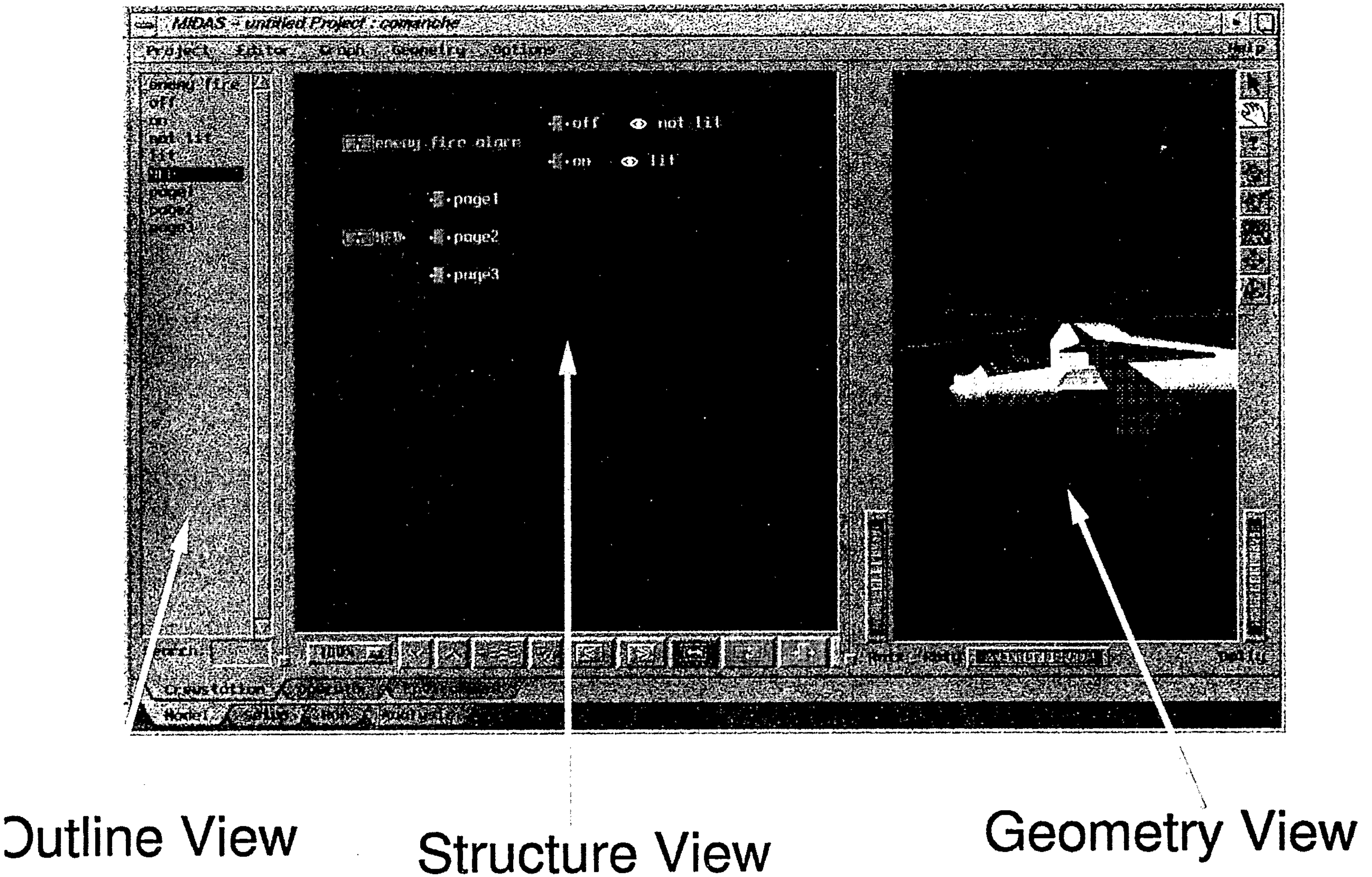


Data Analysis can provide the summarized data for task agenda



MIDAS Redesign Milestone 3

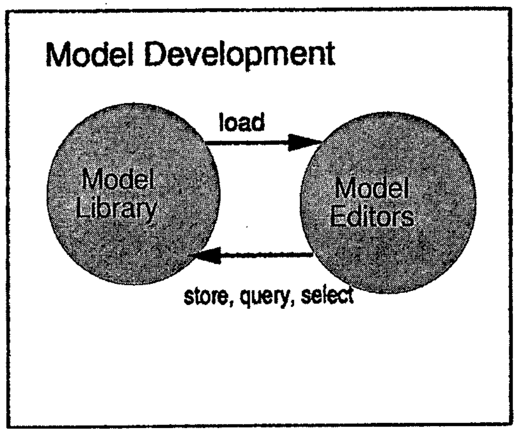
Crewstation Editor

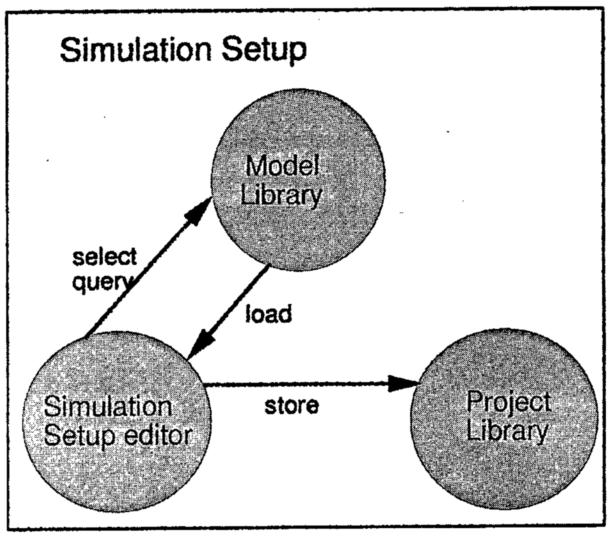


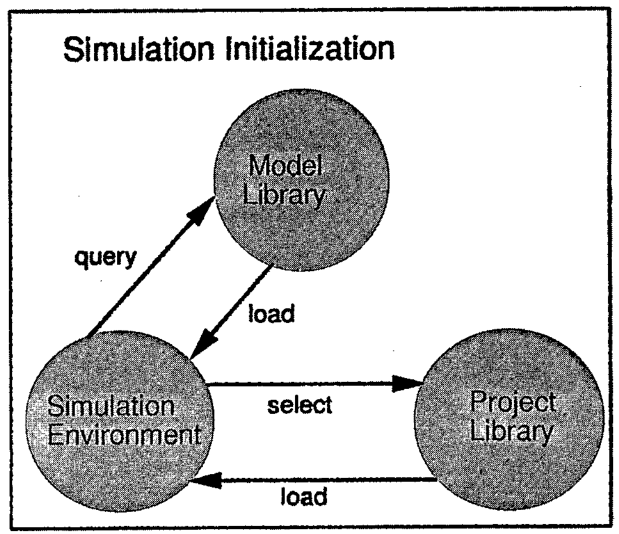
BACK NEXT

MIDAS Redesign Milestone 3

Activities supported by Provisioning





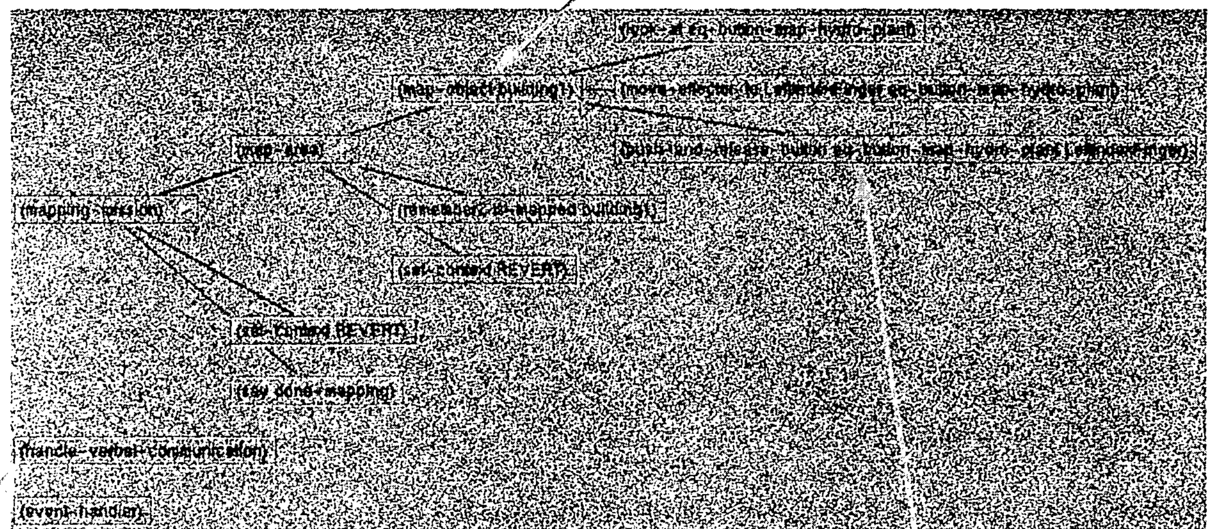


MIDAS Redesign Milestone 3

Task Agenda

Stores instantiated RAPs as 'goals' – complete with current subtask networks; object bindings; control flags; history, etc.

The operator (copilot) has recognized and is mapping the Hydro–electric plant.



For this agenda, there are only sequential tasks.

Note the Event (SA) monitors waiting to trigger

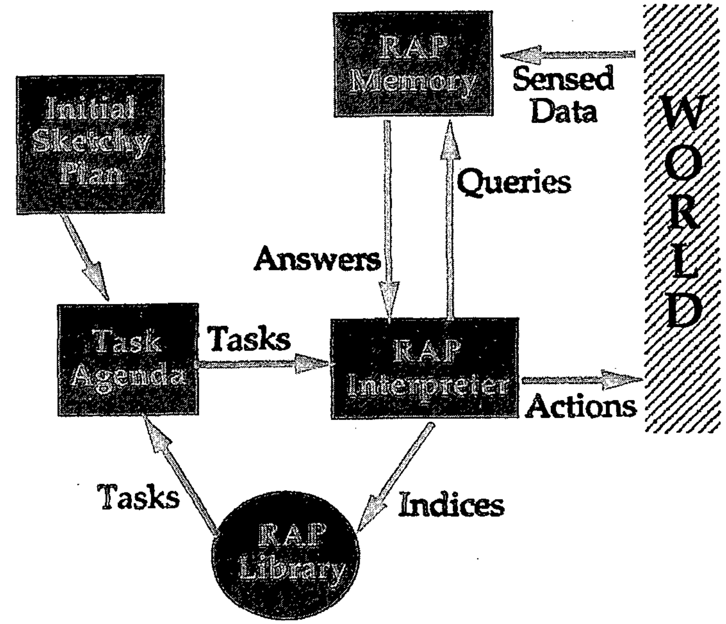
MIDAS Redesign Milestone 3

RAP review

Rap
- Index (*name arg1 .. argn*)
- Succeed (*expr*)
- Monitor–State (*expr*)

Method
- Context–1 (*expr*)
  - TaskNet (*net of goals*)
- ⋮
- Context–*n* (*expr*)
  - TaskNet (*net of goals*)

Method
- Context–1 (*expr*)
  - TaskNet (*net of goals*)
- ⋮
- Context–*n* (*expr*)
  - TaskNet (*net of goals*)

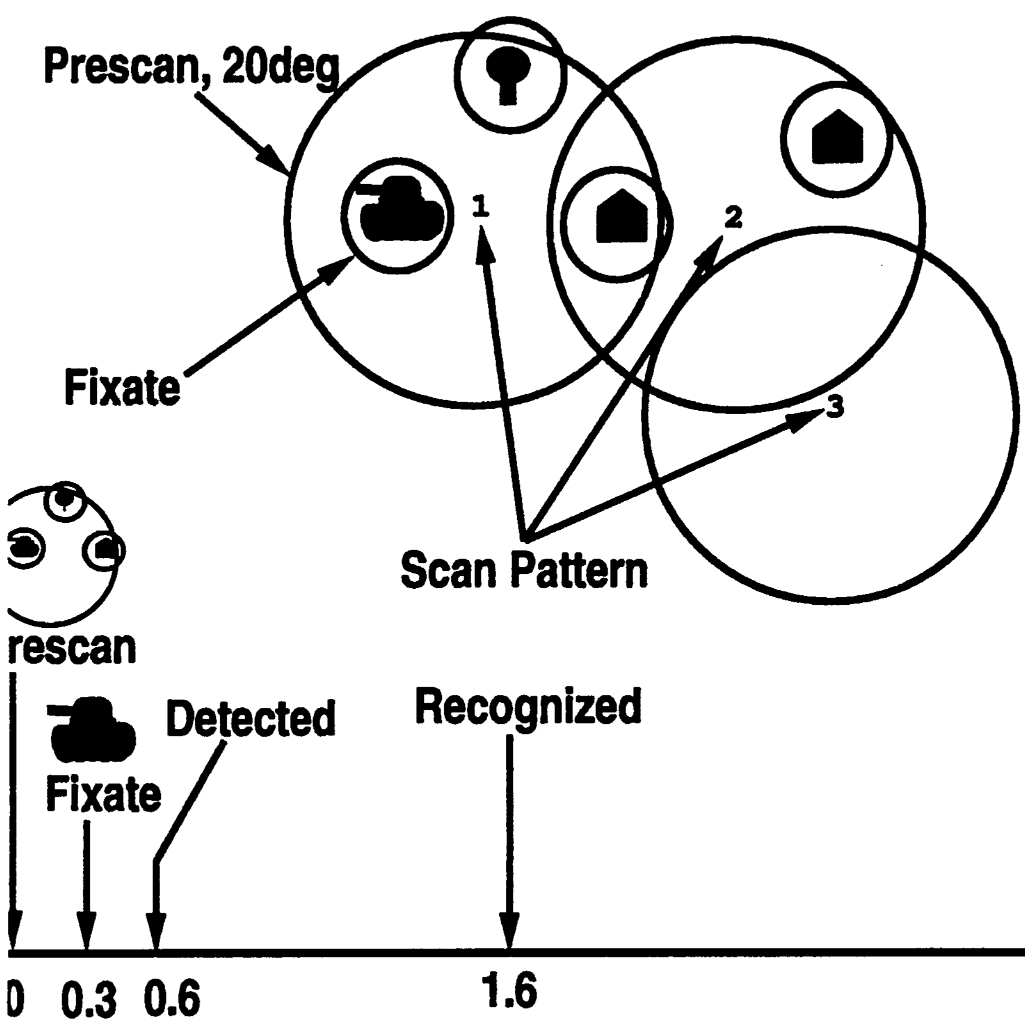



```
(rap
 (index        (mapping-mission))
 (succeed      (and (speech pilot acknowledged)
                    (complete scan)))
 (monitor-state (speech pilot start-mapping))
 (method
  (task-net
   (t1 (begin-scan-area)       (for t2))
   (t2 (set-context scanning)  (for t3))
   (t3 (map-area)              (for t4))
   (t4 (set-context REVERT)    (for t5))
   (t5 (say done-mapping)))))
```

ACK NEXT

MIDAS Redesign Milestone 3

Vision Model: Exterior Scan

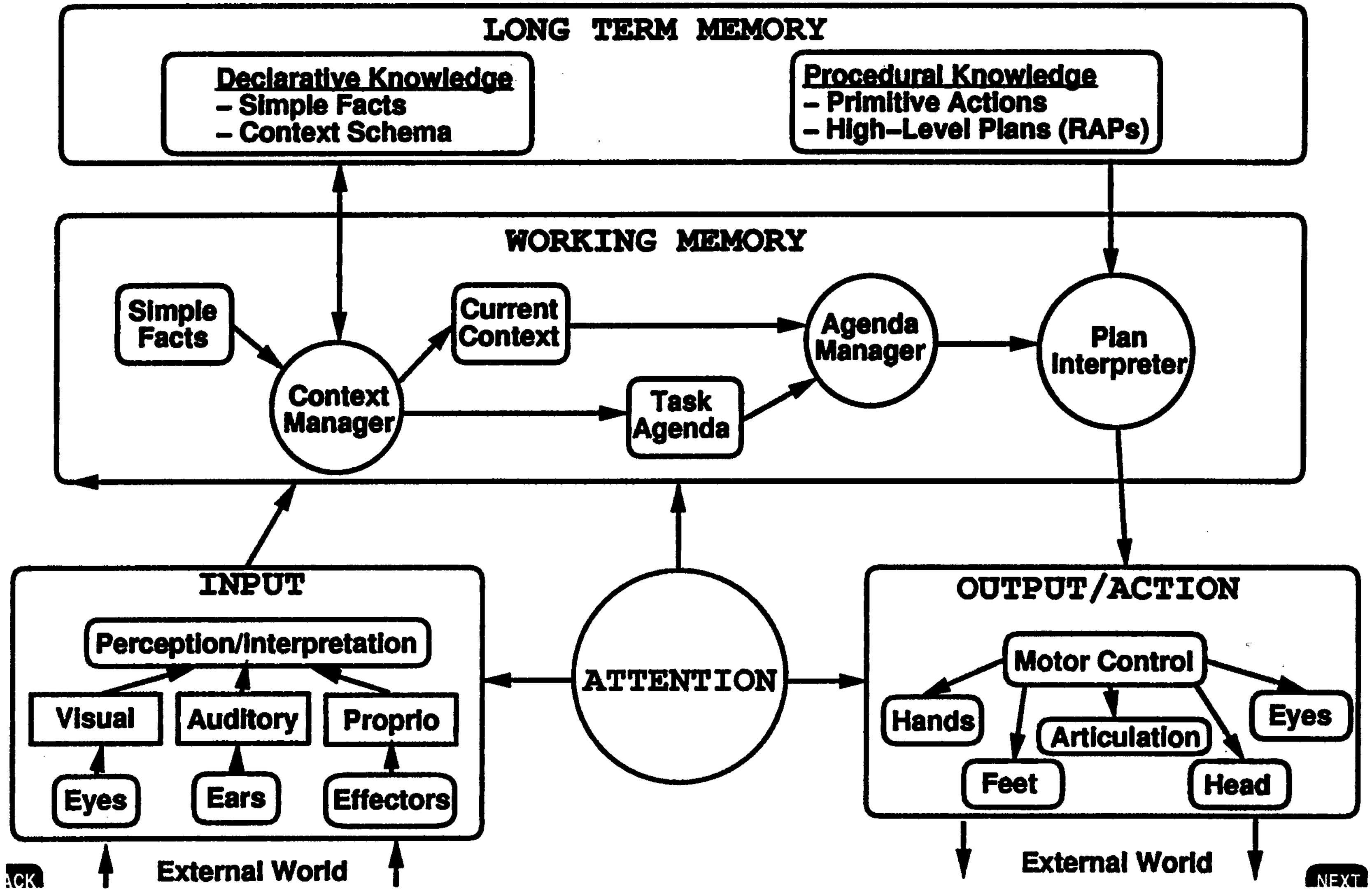


States	Attributes
Detected	size, contrast, movement, location
Recognized	class, shape, motion
Identified	texture, color, transparency ID

ACK NEXT

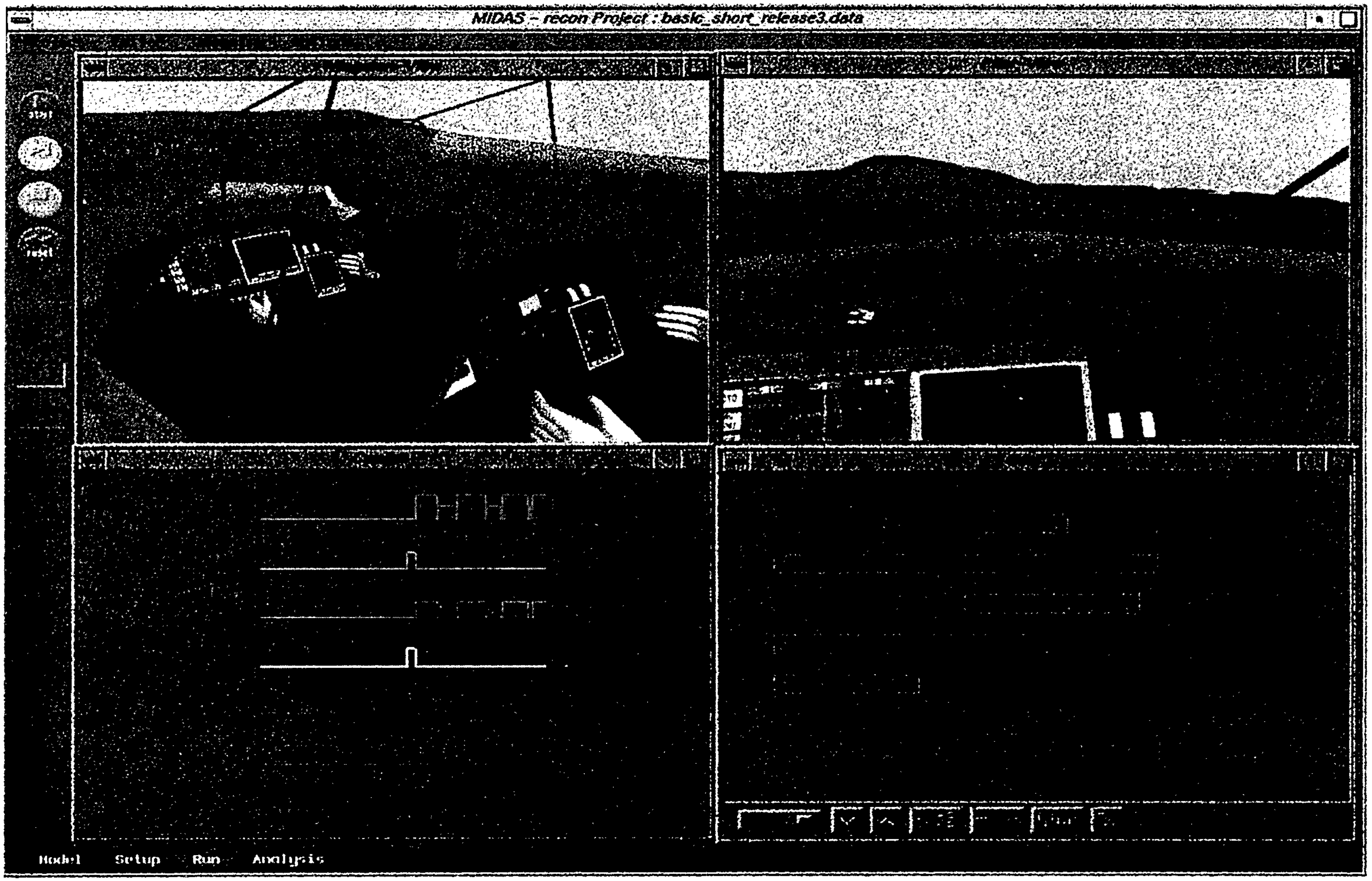
MIDAS Redesign Milestone 3

Human Performance Model: Overview



MIDAS Redesign Milestone 3

Basic Scenario



MIDAS Redesign Milestone 3

Scenario mission

Setting: Takes place aboard a helicopter with two crewstations
Copilot seated in front, pilot in the back
The terrain contains hills and the following features:
grove, sewage plant, blast furnace, hydro–electric plant, tank

Crewstation equipment:

pilot:	wingman signal, enemy fire signal
copilot:	6 buttons (4 mapping, 2 mfd), 1 mfd

Mission:

- Pilot flies route
- Pilot receives signal from wingman

→

- Pilot–>Copilot "StartMapping"
- Copilot scans area, mapping objects of interest

→

- Copilot–>Pilot "DoneMapping"
- Pilot–>Copilot "Acknowledged"

Modified Mission:

- Pilot flies route
- Pilot receives signal from wingman

→

- Pilot–>Copilot "StartMapping"
- Copilot begins to scan the area, mapping objects, until...

→

- Pilot receives enemy incoming fire signal
- Pilot–>Copilot "Enemy–Fire!"
- Copilot aborts msn, switches MFD page, then fires weapon

Basic Scenario

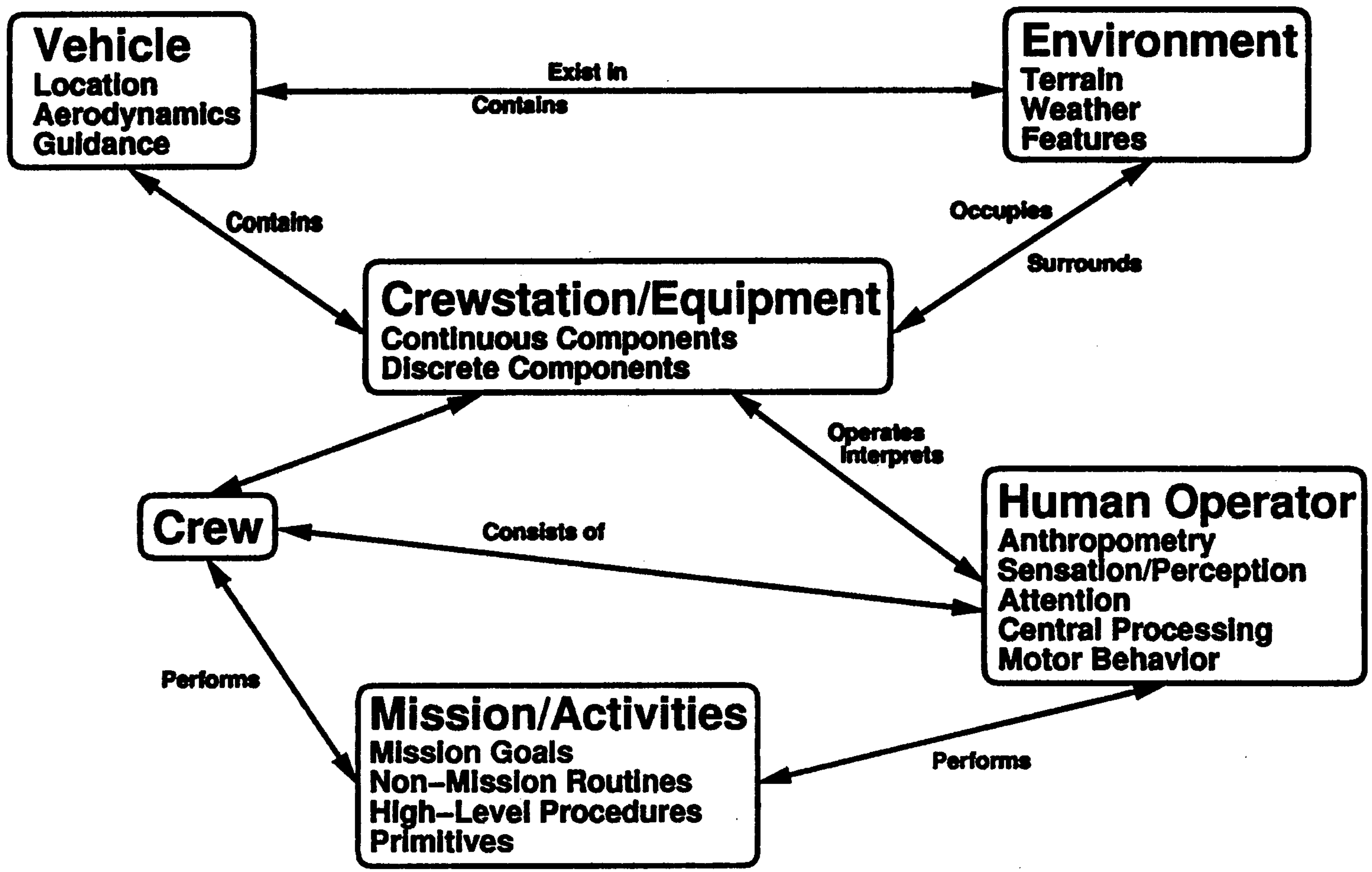
Static scenario demonstrates the following features:

General:
- Terrain with features (trees, buildings, tank)
- Moving vehicle with route & aero model
- Multiple operator and equipment models
- Time scripted events

Operator:
- Perception of auditory signal from equipment
- Taking reading on crewstation equipment
- Scanning exterior of crewstation
- Detection and recognition of objects
- Changing context
- Task interruption & resumption
- Verbal communication between operators

MIDAS Redesign Milestone 3

Recap – Milestone 1: Domain Model



ACK NEXT

APPENDIX B

This appendix contains the matrices used to define the weightings and ratings of the models implemented in the HOMER Expert System. It also contains examples of two completed tables.

B1. Models In HOMER v1.0

Models in HOMER v1.0															
Qualifier	Wght	Variables	MID	PUM	Phrs-2	IPME	Orcl	HOS	Tawl Toss	Win-Crw	W/ Indx	OCM	Pro Cru	Jack	FAIT
My primary	5	crew complement	1	2	-20	2	-20	1	2	4	2	-20	-20	2	1
interests are:	5	team interactions	3	3	-20	3	-20	2	1	2	-20	-20	1	1	1
	5	display format & dynamics	3	1	1	2	3	2	1	2	2	-20	1	2	1
	5	control design & dynamics	3	1	2	3	-20	2	2	1	2	4	3	2	1
	5	automation	4	2	1	3	-20	2	3	4	1	1	2	-20	3
	5	procedures	4	2	1	3	-20	2	1	3	1	1	2	1	-20
	5	workspace geometry & layout	4	1	-20	4	-20	3	-20	3	3	-20	-20	4	1
	5	communications	2	2	1	2	-20	1	1	2	1	-20	1	-20	2
	5	environmental stressors	2	3	2	3	2	2	-20	2	-20	1	1	1	2
Design phase(s)	2	conceptual design	2	3	1	3	2	3	2	1	3	1	1	2	4
I will analyse are:	2	feasibility; dem/val	3	3	2	3	2	3	3	4	3	3	3	3	2
	2	system development	3	2	3	3	2	3	3	4	3	3	3	3	-2
	2	test & evaluation	3	1	2	2	4	2	2	2	2	3	3	4	2
The eqpt/sys :	2	off the shelf	2	3	2	3	4	3	3	3	3	3	3	3	2
I will analyse is	2	mod of existing sys	4	3	4	3	4	3	4	4	4	4	4	4	3
	2	completely new system	4	3	2	3	3	3	2	1	2	2	2	4	2
Complexity of sys	1	simple device	4	4	0	4	4	4	4	4	4	0	2	2	0
I will analyse is:	1	complete complex sys	4	4	4	4	2	4	4	4	0	4	0	4	0
The crew I plan to	3	single operator	4	4	4	4	4	4	4	4	4	4	4	4	4
analyse is:	3	2 or more operators	4	2	2	3	2	2	2	4	3	-4	2	3	2
Max time avail	4	days	-4	-2	-4	-4	-4	-4	-4	-4	-4	-4	-4	-2	0
for completing	4	weeks	-4	-1	-1	-2	0	-2	-2	-2	-1	-1	-4	-1	0
analysis	4	months	0	0	0	0	0	0	0	0	0	0	0	0	0

is:															
Funds available	4	$0-5000	0	-4	-4	0	0	-4	0	0	0	0	-4	-4	0	
for software	4	$5000-50,000	0	0	0	0	0	0	0	0	0	0	-4	-4	0	
purchase are:	4	>$50,000	0	0	0	0	0	0	0	0	0	0	0	0	0	
I do am NOT	4	IBM-type PC (Windows)	0	0	-4	0	0	-4	0	-4	-4	0	0	0	0	
willing to use an	4	PC or Sun (with UNIX)	0	0	0	-4	0	0	0	0	0	0	-4	-4	0	
	4	SGI	-4	0	0	-4	0	0	0	0	0	0	0	0	0	
	4	Macintosh	0	-4	0	0	0	0	0	0	0	0	0	0	0	
	4	any computer	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	0	
My approximate	4	160-640 man hours	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	0	
personnel budget is	4	640-2000 man hours	-3	-1	-1	-2	0	-2	-2	-2	-2	-2	-3	0	0	
	4	> 2000 man hours	0	0	0	0	0	0	0	0	0	0	0	0	0	
Personnel skills	2	subject matter experts	0	0	0	0	0	0	0	0	0	-4	0	-4	-4	
include:	2	human factors experts	0	0	-4	0	0	0	-4	0	-4	-4	-4	-4	-4	
(- if don't have)	2	computer programmers	0	-4	-4	0	-4	-4	-4	-4	-4	0	0	-4	-4	
(0 if have)	2	modeller/systems analyst	0	-4	-4	0	-4	-4	-4	-4	-4	0	0	-4	-4	
Available data	3	timelines	0	-4	0	0	0	0	-4	0	-4	-4	0	0	0	
include:	3	task network	-4	-4	-4	-4	0	-4	0	-4	0	0	0	0	0	
(- if don't have)	3	human, sys, env parameters	-4	0	-4	0	-4	0	-4	0	-4	-4	0	0	0	
(0 if have)	3	analysis of similar system	-2	0	0	-2	-2	0	-2	-2	0	-2	0	0	0	
	3	model of relevant dynamics	-4	0	0	-4	0	0	0	0	0	-4	0	0	0	
Represent wkld	2	mission duration	4	0	0	4	4	4	0	4	0	4	4	4	0	
peaks by:	2	errors	4	4	0	4	4	4	4	4	4	0	0	0	0	
It is important that	4	vehicle control model	4	-4	-4	4	-4	4	-4	-4	-4	4	4	-4	-4	
the model supports:	4	cockpit layout	4	4	-4	4	-4	4	-4	4	-4	-4	4	4	-4	
	4	state transitions	4	-4	-4	4	-4	4	4	4	-4	-4	4	-4	4	
	4	system/automation logic	4	-4	-4	4	-4	-4	-4	4	-4	4	4	-4	-4	

	4	physical sim of workspace	4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	4	-4	
	4	physical simulation of scene	4	4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	
Model runs in:	5	real time	-4	4	-4	4	-4	-4	-4	-4	-4	-4	-4	-4	-4	
	5	faster than real-time	4	-4	-4	4	4	4	-4	4	-4	-4	4	-4	-4	
For decisions, the model must:	3	emulate dec process/generate decisions	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	
	3	generate dec by following user-spec rules	4	-4	-4	4	-4	4	4	4	-4	-4	-4	-4	-4	
	3	introduce user-spec dec at user-spec points	4	-4	-4	4	-4	4	4	4	-4	4	4	-4	-4	
For errors, the model must:	3	generate reasonable errors at likely points	-4	-4	4	4	-4	-4	-4	4	-4	4	4	-4	4	
	3	insert user-specified errors at likely points	-4	-4	4	4	-4	-4	-4	4	-4	4	4	-4	4	
	3	insert user-specified errors at user-specified points	4	-4	-4	4	-4	4	4	4	-4	4	4	-4	-4	
The model must	5	response times	4	-4	-4	4	4	4	4	4	-4	-4	4	-4	-4	
output:	5	accuracy estimates	4	-4	-4	4	4	4	4	4	-4	4	4	-4	-4	
	5	crew workload estimates	4	-4	-4	4	-4	-4	4	4	4	-4	4	-4	-4	
	5	task list	4	-4	-4	4	-4	4	4	4	-4	-4	-4	-4	-4	
	5	task network	4	-4	-4	4	-4	4	-4	4	-4	-4	-4	-4	-4	
	5	procedure list	4	-4	-4	-4	-4	-4	4	4	-4	-4	4	-4	-4	
	5	timeline	4	4	-4	4	-4	4	4	4	4	-4	4	-4	-4	
	5	function/task allocation	4	4	4	4	4	4	4	4	4	4	4	-4	4	
	5	biomechanical measures	4	-4	-4	4	-4	-4	-4	-4	-4	-4	-4	4	-4	
	5	fit, reach, visual envelopes	4	-4	-4	4	4	-4	-4	-4	-4	-4	-4	4	-4	
	5	training requirements	4	4	4	4	4	4	4	4	-4	-4	-4	-4	4	
	5	selection requirements	-4	-4	-4	-4	4	-4	-4	4	-4	-4	-4	4	4	
	5	estimate of sys	4	4	4	4	4	4	4	4	4	4	4	4	4	

		effectiveness													
	5	maintainability	4	4	4	4	-4	4	4	4	4	-4	-4	4	-4	
The output must be:	2	real, absolute values	4	4	4	4	4	4	-4	4	-4	4	4	4	-4	
	2	figures of merit	4	4	4	4	4	4	4	4	4	4	4	4	4	
The model	1	mission, task, crew sum	4	4	4	4	4	4	4	4	-4	4	-4	4	4	
must generate :	1	task/segment summary	4	4	4	4	-4	4	4	4	4	4	-4	-4	-4	
	1	sec by sec events	4	4	-4	4	-4	4	4	4	4	4	4	4	-4	
The model can:	4	Generate a dynamic visualization (animation)	4	4	-4	4	-4	-4	-4	4	-4	-4	-4	4	-4	
Estimate impact on	1	human characteristics	2	1	4	3	4	2	-4	4	-4	4	4	-4	2	
crew/sys perf of:	1	eqpt characteristics	3	1	4	3	-4	2	2	4	1	3	3	-4	2	
	1	environmental factors	2	1	2	3	2	2	-4	2	1	3	3	2	2	
	1	stressors	1	4	2	3	2	3	-4	2	-4	-4	1	-4	2	
															Ratings

B2. HOMER: ASSESSMENT OF MODEL CAPABILITIES - MIDAS

	Topics which can be addressed with the model:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
1	crew complement		¦			
2	team interactions				¦	
3	display format & dynamics				¦	
4	control design & dynamics				¦	
5	automation					¦
6	procedures					¦
7	workspace geometry/layout					¦
8	communications			¦		
9	environmental stressors			¦		
	Design phase(s) it supports:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
10	conceptual design			¦		
11	feasibility; dem/val				¦	
12	system development				¦	
13	test & evaluation				¦	
	Types of equipment or systems it can model/analyse:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
14	off the shelf equipment			¦		
15	modification of existing system					¦
16	completely new system					¦
	System complexity it can accommodate:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
17	simple device					¦
18	complete, complex system					¦
	The number of operators that can be modelled:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
19	single operator					¦
20	2 or more operators					¦
	Minimum time required to develop a model/analysis:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
21	days	¦				
22	weeks		¦			
23	months					¦
	Cost of software:	**Yes**	**No**			
24	$0-5000	¦				
25	$5000-50,000		¦			
26	>$50,000		¦			
	The computer(s) upon which it runs include:	**Yes**	**No**			

27	IBM-type PC (with Windows)		⁝
28	IBM-type PC or Sun (with UNIX)		⁝
29	Silicon Graphics Workstation	⁝	
30	Macintosh		⁝
31	none required		⁝
	Man-hours required to develop a model/analysis:	**Yes**	**No**
32	160-640 man-hours		⁝
33	640-2000 man-hours	⁝	
34	>2000 man-hours	⁝	
	Support personnel required to develop a model/analysis:	**Yes**	**No**
35	subject matter experts	⁝	
36	human factors experts	⁝	
37	computer programmers	⁝	
38	modeller/systems analyst	⁝	
	Data required to develop a model/analysis:	**Yes**	**No**
39	timeline		
40	task network	⁝	
41	human, sys, env parameters	⁝	
42	analysis of similar system	⁝	
43	model of relevant dynamics	⁝	
	Excessive crew workload is represented by a change in:	**Yes**	**No**
44	mission duration	⁝	
45	errors	⁝	
46	indicating overload	⁝	
	The model supports the following:	**Yes**	**No**
47	vehicle control model	⁝	
48	cockpit layout	⁝	
49	state transitions	⁝	
50	system/automation logic	⁝	
51	physical simulation of workspace	⁝	
52	graphic depiction of outside scene	⁝	
	The model can run in:	**Yes**	**No**
53	real time		⁝
54	faster than real time	⁝	
	With respect to decisions, the model:	**Yes**	**No**
55	emulates the decision process & generates decisions		⁝
56	generates decisions by following user-specified rules	⁝	
57	introduces user-specified decisions at user-specified points	⁝	

	With respect to errors, the model:	**Yes**	**No**			
58	generates reasonable errors at likely points		¦			
59	inserts user-specified errors at likely points		¦			
60	inserts user-specified errors at user-specified points	¦				
	The output of the model includes:	**Yes**	**No**			
61	response times	¦				
62	accuracy estimates	¦				
63	crew workload estimates	¦				
64	task list	¦				
65	task network	¦				
66	procedure list	¦				
67	timeline	¦				
68	function/task allocation	¦				
69	biomechanical measures	¦				
70	fit, reach, visual envelopes	¦				
71	training requirements	¦				
72	selection requirements		¦			
73	estimate of system effectiveness	¦				
74	maintainability	¦				
	The output of the model is in the form of:	**Yes**	**No**			
75	absolute values (e.g., RT, RMSe)	¦				
76	figures of merit (e.g., % change)	¦				
	The model can produce:	**Yes**	**No**			
77	summaries by mission, task, crew	¦				
78	summaries by segment	¦				
79	second by second events	¦				
	The model can:	**Yes**	**No**			
80	Generate dynamic visualization (animation)	¦				
	It can estimate the impact on crew/system performance of:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
81	human characteristics			¦		
82	equipment characteristics				¦	
83	environmental factors			¦		
84	physical and emotional stressors		¦			

B3. HOMER: ASSESSMENT OF MODEL CAPABILITIES - ORACLE

	Topics which can be addressed with the model:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
1	crew complement	¦				
2	team interactions	¦				
3	display format & dynamics				¦	
4	control design & dynamics	¦				
5	automation	¦				
6	procedures	¦				
7	workspace geometry/layout	¦				
8	communications	¦				
9	environmental stressors			¦		
	Design phase(s) it supports:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
10	conceptual design			¦		
11	feasibility; dem/val			¦		
12	system development			¦		
13	test & evaluation					¦
	Types of equipment or systems it can model/analyse:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
14	off the shelf equipment					¦
15	modification of existing system					¦
16	completely new system				¦	
	System complexity it can accommodate:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
17	simple device					¦
18	complete, complex system			¦		
	The number of operators that can be modelled:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
19	single operator					¦
20	2 or more operators			¦		
	Minimum time required to develop a model/analysis:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
21	days	¦				
22	weeks					¦
23	months					¦
	Cost of software:	**Yes**	**No**			
24	$0-5000	¦				
25	$5000-50,000		¦			
26	>$50,000		¦			
	The computer(s) upon which runs include:	**Yes**	**No**			
27	IBM-type PC (with Windows)					

28	IBM-type PC or Sun (with UNIX)		
29	Silicon Graphics Workstation		
30	Macintosh		
31	none required		
	Man-hours required to develop a model/analysis:	**Yes**	**No**
32	160-640 man-hours		¦
33	640-2000 man-hours	¦	
34	>2000 man-hours	¦	
	Support personnel required to develop a model/analysis:	**Yes**	**No**
35	subject matter experts		¦
36	human factors experts		¦
37	computer programmers	¦	
38	modeller/systems analyst	¦	
	Data required to develop a model/analysis:	**Yes**	**No**
39	timeline		¦
40	task network		¦
41	human, sys, env parameters	¦	
42	analysis of similar system	¦	
43	model of relevant dynamics		¦
	Excessive crew workload is represented by a change in:	**Yes**	**No**
44	mission duration	¦	
45	errors	¦	
46	indicating overload		¦
	The model supports the following:	**Yes**	**No**
47	vehicle control model		¦
48	cockpit layout		¦
49	state transitions		¦
50	system/automation logic		¦
51	physical simulation of workspace		¦
52	graphic depiction of outside scene		¦
	The model can run in:	**Yes**	**No**
53	real time		¦
54	faster than real time	¦	
	With respect to decisions, the model:	**Yes**	**No**
55	emulates the decision process & generates decisions		¦
56	generates decisions by following user-specified rules		¦
57	introduces user-specified decisions at user-specified points		¦

	With respect to errors, the model:	**Yes**	**No**			
58	generates reasonable errors at likely points		¦			
59	inserts user-specified errors at likely points		¦			
60	inserts user-specified errors at user-specified points		¦			
	The output of the model includes:	**Yes**	**No**			
61	response times	¦				
62	accuracy estimates	¦				
63	crew workload estimates		¦			
64	task list		¦			
65	task network		¦			
66	procedure list		¦			
67	timeline		¦			
68	function/task allocation	¦				
69	biomechanical measures		¦			
70	fit, reach, visual envelopes	¦				
71	training requirements	¦				
72	selection requirements	¦				
73	estimate of system effectiveness	¦				
74	maintainability		¦			
	The output of the model is in the form of:	**Yes**	**No**			
75	absolute values (e.g., RT, RMSe)	¦				
76	figures of merit (e.g., % change)	¦				
	The model can produce:	**Yes**	**No**			
77	summaries by mission, task, crew	¦				
78	summaries by segment		¦			
79	second by second events		¦			
	The model can:	**Yes**	**No**			
80	Generate dynamic visualization (animation)		¦			
	It can estimate the impact on crew/system performance of:	Can't generate or support..	can do it with difficulty	can do it adequately	can do it well	can do it extremely well
81	human characteristics					¦
82	equipment characteristics	¦				
83	environmental factors			¦		
84	physical and emotional stressors			¦		

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-AR-356	ISBN 92-836-1077-6	UNCLASSIFIED/ UNLIMITED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
BP 25, 7 rue Ancelle, F-92201 Neuilly-sur-Seine Cedex, France

6. Title
A Designer's Guide to Human Performance Modelling

7. Presented at/sponsored by
The request of the Aerospace Medical Panel.

8. Author(s)/Editor(s)	9. Date
Multiple	December 1998

10. Author's/Editor's Address	11. Pages
Multiple	168

12. Distribution Statement There are no restrictions on the distribution of this document. Information about the availability of this and other AGARD unclassified publications is given on the back cover.

13. Keywords/Descriptors

Human factors engineering	Operations research	Control theory
Performance	Case studies	Biomechanics
Models	Reliability	Senses
Man machine systems	Situational awareness	Cognition
Man computer interface	Workloads	Perception
Systems engineering	Predictions	Visual perception
Operators (personnel)	Anthropometry	Errors
Design	Task analysis	Workplace layout
Expert computer systems	Automatic control	Decision making

14. Abstract

Working Group 22 was convened in 1995, jointly sponsored by the Aerospace Medical Panel and the Flight Vehicle Panel to investigate the use of Human Performance Models within the specification, procurement, design, qualification and certification of military systems.

In particular the group focused on the selection, application and use of HPMs by the system designer. An expert system approach was selected to ensure that the designer considered all the relevant factors when selecting a new model or tool. This was implemented using a commercially available expert system shell. The user is asked to select options that most closely describe his resources and requirements and the Human Operator Modelling Expert Review (HOMER) then rank-orders the HPMs in its database and suggests the most appropriate model.

The group carried out some walkthroughs of existing models/tools to demonstrate typical uses in the analysis of specific issues. These are included as case studies. These were included to give potential users some insight into the ease or complexity of use in order to evaluate the required aspect of human performance.

In addition the group also considered the model developer community by examining the limitations of existing models, commercial implications and usability issues in order to guide any future development.

AGARD

NATO - OTAN

BP 25 • 7 RUE ANCELLE

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Télécopie 0(1)55.61.22.99 • Télex 610 176

DIFFUSION DES PUBLICATIONS

AGARD NON CLASSIFIEES

L'AGARD détient un stock limité de certaines de ses publications récentes. Celles-ci pourront éventuellement être obtenus sous forme de copie papier. Pour de plus amples renseignements concernant l'achat de ces ouvrages, adressez-vous à l'AGARD par lettre ou par télécopie à l'adresse indiquée ci-dessus. *Veuillez ne pas téléphoner.*

Des exemplaires supplémentaires peuvent parfois être obtenus auprès des centres de diffusion nationaux indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de l'AGARD, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus sur la liste d'envoi de l'un de ces centres.

Les publications de l'AGARD sont en vente auprès des agences de vente indiquées ci-dessous, sous forme de photocopie ou de microfiche. Certains originaux peuvent également être obtenus auprès de CASI.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

BELGIQUE
Coordonnateur AGARD - VSL
Etat-major de la Force aérienne
Quartier Reine Elisabeth
Rue d'Evere, B-1140 Bruxelles

CANADA
Directeur - Gestion de l'information
(Recherche et développement) - DRDGI 3
Ministère de la Défense nationale
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

ESPAGNE
INTA (AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

ETATS-UNIS
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

FRANCE
O.N.E.R.A. (Direction)
29, Avenue de la Division Leclerc
92322 Châtillon Cedex

GRECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ISLANDE
Director of Aviation
c/o Flugrad
Reykjavik

ITALIE
Aeronautica Militare
Ufficio Stralcio AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PAYS-BAS
Netherlands Delegation to AGARD
National Aerospace Laboratory NLR
P.O. Box 90502
1006 BM Amsterdam

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

ROYAUME-UNI
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

TURQUIE
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

AGENCES DE VENTE

NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320
Etats-Unis

The British Library Document Supply Division
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
Royaume-Uni

Les demandes de microfiches ou de photocopies de documents AGARD (y compris les demandes faites auprès du CASI) doivent comporter la dénomination AGARD, ainsi que le numéro de série d'AGARD (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Veuiller noter qu'il y a lieu de spécifier AGARD-R-nnn et AGARD-AR-nnn lors de la commande des rapports AGARD et des rapports consultatifs AGARD respectivement. Des références bibliographiques complètes ainsi que des résumés des publications AGARD figurent dans les journaux suivants:

Scientific and Technical Aerospace Reports (STAR)
STAR peut être consulté en ligne au localisateur de ressources uniformes (URL) suivant:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR est édité par CASI dans le cadre du programme NASA d'information scientifique et technique (STI)
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
Etats-Unis

Government Reports Announcements & Index (GRA&I)
publié par le National Technical Information Service
Springfield
Virginia 2216
Etats-Unis
(accessible également en mode interactif dans la base de données bibliographiques en ligne du NTIS, et sur CD-ROM)

Imprimé par le Groupe Communication Canada Inc.
(membre de la Corporation St-Joseph)
45, boul. Sacré-Cœur, Hull (Québec), Canada K1A 0S7

AGARD

NATO - OTAN

BP 25 • 7 RUE ANCELLE

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Telefax 0(1)55.61.22.99 • Telex 610 176

DISTRIBUTION OF UNCLASSIFIED

AGARD PUBLICATIONS

AGARD holds limited quantities of some of its recent publications, and these may be available for purchase in hard copy form. For more information, write or send a telefax to the address given above. *Please do not telephone.*

Further copies are sometimes available from the National Distribution Centres listed below. If you wish to receive all AGARD publications, or just those relating to one or more specific AGARD Panels, they may be willing to include you (or your organisation) in their distribution.

AGARD publications may be purchased from the Sales Agencies listed below, in photocopy or microfiche form. Original copies of some publications may be available from CASI.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD - VSL
Etat-major de la Force aérienne
Quartier Reine Elisabeth
Rue d'Evere, B-1140 Bruxelles

CANADA
Director Research & Development
Information Management - DRDIM 3
Dept of National Defence
Ottawa, Ontario K1A 0K2

DENMARK
Danish Defence Research Establishment
Ryvangs Allé 1
P.O. Box 2715
DK-2100 Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92322 Châtillon Cedex

GERMANY
Fachinformationszentrum Karlsruhe
D-76344 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force
Air War College
Scientific and Technical Library
Dekelia Air Force Base
Dekelia, Athens TGA 1010

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio Stralcio AGARD
Aeroporto Pratica di Mare
00040 Pomezia (Roma)

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 90502
1006 BM Amsterdam

NORWAY
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Estado Maior da Força Aérea
SDFA - Centro de Documentação
Alfragide
P-2720 Amadora

SPAIN
INTA (AGARD Publications)
Carretera de Torrejón a Ajalvir, Pk.4
28850 Torrejón de Ardoz - Madrid

TURKEY
Millî Savunma Başkanliği (MSB)
ARGE Dairesi Başkanliği (MSB)
06650 Bakanliklar - Ankara

UNITED KINGDOM
Defence Research Information Centre
Kentigern House
65 Brown Street
Glasgow G2 8EX

UNITED STATES
NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320

SALES AGENCIES

NASA Center for AeroSpace Information (CASI)
Parkway Center, 7121 Standard Drive
Hanover, MD 21076-1320
United States

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
United Kingdom

Requests for microfiches or photocopies of AGARD documents (including requests to CASI) should include the word 'AGARD' and the AGARD serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Note that AGARD Reports and Advisory Reports should be specified as AGARD-R-nnn and AGARD-AR-nnn, respectively. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
STAR is available on-line at the following uniform resource locator:
http://www.sti.nasa.gov/Pubs/star/Star.html
STAR is published by CASI for the NASA Scientific and Technical Information (STI) Program
STI Program Office, MS 157A
NASA Langley Research Center
Hampton, Virginia 23681-0001
United States

Government Reports Announcements & Index (GRA&I)
published by the National Technical Information Service
Springfield
Virginia 22161
United States
(also available online in the NTIS Bibliographic Database or on CD-ROM)

Printed by Canada Communication Group Inc.
(A St. Joseph Corporation Company)
45 Sacré-Cœur Blvd., Hull (Québec), Canada K1A 0S7

ISBN 92-836-1077-6